

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 05:45:20 ; Search time 3396 Seconds
(without alignments)
10872.460 Million cell updates/sec

Title: US-09-441-857-11
Perfect score: 762

Sequence: 1 atgtccggggcgagactt.....tgcgtcgtctctgtcgagaga 762

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_com.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_scs.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	761	99.9	765	9	CR533469
2	761	99.9	840	9	AF309646 Homo sapi
3	761	99.9	3084	6	AX086357 Sequence
4	759.4	99.7	198961	9	AC079776 Homo sapi
5	759.4	99.7	349981	6	AX647880 Sequence
6	759.4	99.7	349989	6	AX647177 Sequence
7	757.8	99.4	165411	9	AC018865 Homo sapi
8	755.6	99.2	195280	9	AC073869 Homo sapi
9	744.4	97.7	242865	2	AC120839 Pan trogl
10	741.2	97.3	184706	9	AC120781 Pan trogl
11	684	89.8	3070	9	BC003617 Homo sapi
12	684	89.8	3100	9	BC068486 Homo sapi
13	684	89.8	3120	9	AC057157 Homo sapi
14	639.2	83.9	3745	6	AR447780 Sequence
15	638.2	83.8	1770	10	BC019118 Mus muscu
16	609.2	79.9	739	9	AF119836 Homo sapi
17	591.8	77.7	683	9	AF130122 Homo sapi
18	587.6	77.1	627	9	AF498941 Homo sapi
19	586	76.9	627	9	AF198616 Homo sapi

20	586	76.9	627	9	CR541984
21	585.4	76.8	1950	10	AB041575 Mus muscu
22	559.6	73.4	740	6	AX771444 Sequence
23	559.6	73.4	740	9	HUMRAB6A
24	554.8	72.8	1170	5	CR390639
25	542.8	71.2	627	9	AF130986
26	542.8	71.2	627	9	AF498939
27	532.6	69.9	844	5	AF540409
28	526.2	69.1	2527	10	BC026915
29	526	69.0	573	9	AF124200
30	477.8	62.7	172962	10	AL713865
31	471.6	61.9	2099	5	BC046683
32	429	56.3	1461	5	BC064230
33	401	52.6	2456	6	AX713746
34	401	52.6	2456	9	AX055504
35	384.8	50.5	1261	9	BC002510
36	384.4	50.4	627	9	BT007263
37	384.4	50.4	2498	5	BC058059
38	383.2	50.3	1266	9	AF166492
39	383.2	50.3	1589	6	BD131169
40	382.8	50.2	627	9	AF498940
41	381.6	50.1	2005	5	BC080215
42	381.6	50.1	5395	9	BC078662
43	375.4	49.3	4768	10	BC060618
44	369	48.4	1652	5	BC074238
45	352.2	46.2	3826	6	AX647191

ALIGNMENTS

RESULT 1

CR533469

LOCUS

DEFINITION

Homo sapiens full open reading frame cDNA clone RZPD0834C0517D for gene RAB6C, RAB6C, member RAS oncogene family; complete cds, incl. stopcodon.

ACCESSION

CR533469.1 GI:49065363

KEYWORDS

Full ORF shuttle clone, Gateway(TM), complete cds.

SOURCE

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 765)

Ebert L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.

Cloning of human full open reading frames in Gateway(TM) system

entry vector (pDONR201)

Unpublished

2 (bases 1 to 765)

Ebert L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.

Direct Submission

Submitted (21-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer

Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,

Germany

RZPD; RZPD0834C0517D, ORFNO 2784

www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPD0834C0517D RZPDLIB;

Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.

834

www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834

www.rzpd.de/products/orfclones/

Contact: Inge Arlart

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 100

Fax: +49 30 32639 111

www.rzpd.de

This clone is available from RZPD;

contact RZPD (customer.service@rzpd.de) for further information.

This CDS clone is a part of a collection of human full ORF clones

generated by RZPD.

This CDS has been inserted into pDONR201 via a BP Clonase(TM)

The CDS has been inserted into pDONR201 via a BP Clonase(TM)

repeat_region	/rpt_family="MaLR"
18790..19092	
repeat_region	/rpt_family="Li"
19093..19385	
repeat_region	/rpt_family="Alu"
19386..19447	
repeat_region	/rpt_family="Li"
19464..19575	
repeat_region	/rpt_family="MaLR"
20644..21086	
repeat_region	/rpt_family="Li"
21087..21384	
repeat_region	/rpt_family="Alu"
21385..21661	
repeat_region	/rpt_family="Li"
22597..22955	
repeat_region	/rpt_family="Li"
23920..24230	
repeat_region	/rpt_family="Alu"
25572..25876	
repeat_region	/rpt_family="Alu"
25885..26189	
repeat_region	/rpt_family="Alu"
27222..27602	
repeat_region	/rpt_family="MaLR"
29239..29359	
repeat_region	/rpt_family="Li"
31097..31251	
repeat_region	/rpt_family="MIR"
31879..32001	
repeat_region	/rpt_family="Achobo"
Query Match 99.7%; Score 759.4; DB 9; Length 198961;	
Best Local Similarity 99.9%; Pred. No. 5.8e-221;	
Matches 760; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 1	ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGAAATCAAGCTGGTGTCTCTGGGG 60
DB 152358	ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGAAATCAAGCTGGTGTCTCTGGGG 152417
QY 61	GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTCACAC 120
DB 152418	GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTCACAC 152477
QY 121	ACCTATCAGGCAATTAATGGCAATGACTTTTATCAAAATCTATGTACTTGGAGATGGA 180
DB 152478	ACCTATCAGGCAATTAATGGCAATGACTTTTATCAAAATCTATGTACTTGGAGATGGA 152537
QY 181	ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGAAACGTCTCCGTAGCCTCATTC 240
DB 152538	ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGAAACGTCTCCGTAGCCTCATTC 152597
QY 241	AGGTACATCCGTGATCTGCTGACGTGTAGTAGTTTACGATATCACAAATGTTAACTCA 300
DB 152598	AGGTACATCCGTGATCTGCTGACGTGTAGTAGTTTACGATATCACAAATGTTAACTCA 152657
QY 301	TTCCAGCAAACTACAAAGTGGATTTGATGATGTGAGAAACAGAAAGAGGAGTGTATTC 360
DB 152658	TTCCAGCAAACTACAAAGTGGATTTGATGATGTGAGAAACAGAAAGAGGAGTGTATTC 152717
QY 361	ATCAGCTAGTAGGAAATAGAACAGATCTTGTGACAGAGCAAGTGTGAGTGGAGGAG 420
DB 152718	ATCAGCTAGTAGGAAATAGAACAGATCTTGTGACAGAGCAAGTGTGAGTGGAGGAG 152777
QY 421	GGAGAGGAAAGCCAAAGGCTGATGTTAGTTTATTTGAAATAGGCAAAACTGGA 480
DB 152778	GGAGAGGAAAGCCAAAGGCTGATGTTAGTTTATTTGAAATAGGCAAAAGCTGGA 152837
QY 481	TACAATGTAAAGCAGCTTTTCGAGCTGTAGCAGCAGCTTTTCGCGGAATGAAAGCACA 540
DB 152838	TACAATGTAAAGCAGCTTTTCGAGCTGTAGCAGCAGCTTTTCGCGGAATGAAAGCACA 152897
QY 541	CAGGACGGAAGCAGAGACATGATGTGCATATAAACTGGAAAAGCCTCAGGAGCAACA 600

DB 152898	CAGGACGGAAGCAGAGACATGAGTGACATATAAACTGGAAAAGCCTCAGGAGCAACA 152957
QY 601	GTCAGCGAAGGGGGTGTCTCTGCTACTCTCCATGTCTTCAACCTTCTCCTCAGAAG 660
DB 152958	GTCAGCGAAGGGGGTGTCTCTGCTACTCTCCATGTCTTCAACCTTCTCCTCAGAAG 153017
QY 661	CCCCCTTACTCTTTTCATTGACTGTCAGTGTGAATATTGGCTTGAACCTTTTCCCTTCATTA 720
DB 153018	CCCCCTTACTCTTTTCATTGACTGTCAGTGTGAATATTGGCTTGAACCTTTTCCCTTCATTA 153077
QY 721	ATAAGCTTTTGCATTCATTCATTCGCTGCTGCTCTGCTGAG 761
DB 153078	ATAAGCTTTTGCATTCATTCATTCGCTGCTGCTGCTGAG 153118
RESULT 5	
AX647880/c	
LOCUS	AX647880 349981 bp DNA linear PAT 04-MAR-2003
DEFINITION	Sequence 2072 from Patent EP1270724.
ACCESSION	AX647880
VERSION	AX647880.1 GI:28802632
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
AUTHORS	Guanosine triphosphate-binding protein coupled receptors
TITLE	Patent: EP 1270724-A 2072 02-JAN-2003;
JOURNAL	National Institute of Advanced Industrial Science and Technology
	(JP) ; Center for Advanced Science and Technology Incubation, Ltd.
	(JP)
FEATURES	Location/Qualifiers
source	1..349981
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="caxon:9606"
	/note="old seq 1369: original length 744,802 splitted into
	3 new seqs-new seq 1369: from 000.001 to 349,980-new seq
	2072: from 300.001 to 649,980-new seq 2073: from 600.001
	to 744,802"
ORIGIN	
Query Match 99.7%; Score 759.4; DB 6; Length 349981;	
Best Local Similarity 99.9%; Pred. No. 6.3e-221;	
Matches 760; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 1	ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGAAATCAAGCTGGTGTCTCTGGGG 60
DB 45699	ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGAAATCAAGCTGGTGTCTCTGGGG 45640
QY 61	GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTCACAC 120
DB 45639	GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTCACAC 45580
QY 121	ACCTATCAGGCAATTAATGGCAATGACTTTTATCAAAATCTATGTACTTGGAGATGGA 180
DB 45579	ACCTATCAGGCAATTAATGGCAATGACTTTTATCAAAATCTATGTACTTGGAGATGGA 45520
QY 181	ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGAAACGTCTCCGTAGCCTCATTC 240
DB 45519	ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGAAACGTCTCCGTAGCCTCATTC 45460
QY 241	AGGTACATCCGTGATCTGCTGACGTGTAGTAGTTTACGATATCACAAATGTTAACTCA 300
DB 45459	AGGTACATCCGTGATCTGCTGACGTGTAGTAGTTTACGATATCACAAATGTTAACTCA 45400
QY 301	TTCCAGCAAACTACAAAGTGGATTTGATGATGTGAGAAACAGAAAGAGGAGTGTATTC 360
DB 45399	TTCCAGCAAACTACAAAGTGGATTTGATGATGTGAGAAACAGAAAGAGGAGTGTATTC 45340

AUTHORS

TITLE The sequence of Homo sapiens BAC clone RP11-32C20
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 165411)

AUTHORS

TITLE Direct Submission
JOURNAL Submitted (21-DEC-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 165411)

AUTHORS

TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 165411)

AUTHORS

TITLE Direct Submission
JOURNAL Submitted (23-NOV-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Nov 23, 2002 this sequence version replaced gi:22830544.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0032C20

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

There is an unresolved tandem repeat from base 115267 to 124390.

Polymorphisms have been identified between AC018804 and this sequence.

Data from AC018804 was used to finish this clone.

This sequence is the entire insert of the clone.

Location/Qualifiers

1..165411
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

source

```

/chromosome="2"
/map="2"
/clone="RP11-32C20"
/clone_lib="RPCI-11"
487..1684
/rpt_family="L1"
repeat_region
1678..1885
/rpt_family="L1"
repeat_region
1890..1986
/rpt_family="L1"
repeat_region
1987..2289
/rpt_family="Alu"
repeat_region
2294..2380
/rpt_family="GA-rich"
repeat_region
2392..2699
/rpt_family="L1"
repeat_region
2905..3262
/rpt_family="Alu"
repeat_region
3284..3395
/rpt_family="T-rich"
repeat_region
3601..3674
/rpt_family="MER105"
repeat_region
3726..3845
/rpt_family="Alu"
repeat_region
3846..3869
/rpt_family="(TAA)n"
repeat_region
3870..4042
/rpt_family="Alu"
repeat_region
4064..4364
/rpt_family="ERV1"
repeat_region
4692..4980
/rpt_family="Alu"
repeat_region
4981..5001
/rpt_family="AT-rich"
repeat_region
5006..5316
/rpt_family="Alu"
repeat_region
5340..5519
/rpt_family="ERV1"
repeat_region
5563..5604
/rpt_family="AT-rich"
repeat_region
5637..5678
/rpt_family="AT-rich"
repeat_region
5757..5781
/rpt_family="AT-rich"
repeat_region
5782..6005
/rpt_family="Alu"
repeat_region
6067..6146
/rpt_family="MER1_type"
repeat_region
6196..6487
/rpt_family="Alu"
repeat_region
6515..6634
/rpt_family="MER1_type"
repeat_region
6755..7044
/rpt_family="Alu"
repeat_region
7349..7641
/rpt_family="Alu"
repeat_region
7644..7932
/rpt_family="Alu"
repeat_region
7948..8108
/rpt_family="BC200"
repeat_region
8294..8589
/rpt_family="Alu"
repeat_region
8824..9036
/rpt_family="L1"
repeat_region
9037..9342
/rpt_family="Alu"
repeat_region
9343..9404
/rpt_family="L1"
repeat_region
9405..9444
/rpt_family="Alu"
repeat_region
9445..9472
/rpt_family="(TTTTG)n"
repeat_region
9473..9748

```


clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1996) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-433A19. Actual start of this clone is at base position 1 of RP11-109E12; actual end is at base position 195280 of RP11-109E12.

Data from AC022629 was used to finish AC073869.

FEATURES

source

Location/Qualifiers

1..195280

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-109E12"

/clone_lib="RPCI-11"

1..167

/note="match to EST AA904412 (NID:g3039535) ok07e01.s1"

82..96

/note="similar to Homo sapiens EST AV651649

(NID:g9872663)"

298..778

/note="match to EST AW293618 (NID:g6700254)"

408..431

/note="similar to Homo sapiens EST AV651649

(NID:g9872663)"

503..670

/note="match to EST AI352084 (NID:g4089290) qr10b01.x1"

651..1100

/note="match to EST AI138580 (NID:g36444552) qd85b12.x1"

695..710

/note="similar to Homo sapiens EST BE791748

(NID:g10213946)"

1698..2014

/note="similar to Homo sapiens EST N30408 (NID:g1148928)

yy13c09.s1"

2015..2410

/note="similar to Homo sapiens EST AV651649

(NID:g9872663)"

2253..2732

/note="match to EST AW340014 (NID:g6836640) hc91e04.x1"

2338..3119

/note="similar to Homo sapiens EST AU139340

(NID:g1100861)"

2340..2885

/note="similar to Bos taurus EST BE726224 (NID:g10123920)"

3154..3941

misc_feature
/note="similar to Canis familiaris EST BI430421
(NID:gl5207653)"
3236..3645
/note="similar to Mus musculus EST AA617192 (NID:g2504397)
vj79a09.r1"
4375..4376
/note="similar to Homo sapiens EST A1677857 (NID:g4888039)
wd34c04.x1"
4427..4468
/rpt_family="(TTTA)n"
4440..4738
/rpt_family="Alu"
4835..4973
/note="similar to Homo sapiens EST A1677857 (NID:g4888039)
wd34c04.x1"
4892..5127
/note="similar to Homo sapiens EST BF511460
(NID:gl1594758)"
4899..5106
/note="match to EST AW804021 (NID:g7855891)"
4974..5130
/note="similar to Homo sapiens EST A1677857 (NID:g4888039)
wd34c04.x1"
5205..5387
/note="similar to Homo sapiens EST BF511460
(NID:gl1594758)"
5368..5470
/rpt_family="(CGG)n"
6103..6166
/rpt_family="(GGGA)n"
6384..6390
/note="match to EST AL042663 (NID:g5935522)"
6456..6485
/rpt_family="(CA)n"
6550..6732
/rpt_family="MIR"
7058..7505
/note="similar to Homo sapiens EST BE646288
(NID:g9970599)"
7216..7316
/rpt_family="MIR"
7327..7378
/rpt_family="MER1_type"
8007..8043
/rpt_family="AT_rich"
8009..8118
/rpt_family="Alu"
8123..8336
/rpt_family="L1"
8338..8442
/rpt_family="L1"
8443..8752
/rpt_family="Alu"
8753..9325
/rpt_family="L1"
9488..9910
/note="match to EST BE973696 (NID:g10587032)"
9763..9944
/note="similar to Homo sapiens EST BG739917
(NID:g14050570)"
9763..9944
/note="similar to Homo sapiens EST R18754 (NID:g772364)
vg17d08.r1"
9767..9981
/note="match to EST BE859007 (NID:g10374638)"
10270..10476
/rpt_family="MIR"
10803..10855
/rpt_family="AT_rich"
11465..11542
/note="match to EST BE973696 (NID:g10587032)"
11771..12131
/rpt_family="L2"

```
misc_feature 11798..11948
/note="match to EST BE973696 (NID:g10587032)"
misc_feature 12659..12660
/note="similar to Homo sapiens EST AW891034
(NID:g8055239)"
misc_feature 12666..12684

Query Match 99.2%; Score 755.6; DB 9; Length 195280;
Best Local Similarity 99.5%; Pred. No. 8.5e-220;
Matches 758; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGTCGGGGGGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTGTCTCTGGGG 60
Db 37987 ATGTCGGGGGGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTGTCTCTGGGG 37928

Qy 61 GAGCAAGCGTTCGAAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGAACAAC 120
Db 37927 GAGCAAGCGTTCGAAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGAACAAC 37868

Qy 121 ACCTATCAGGCAATATTTGGCATTGACTTTTATCAAAAATATGTACTTCGGAGATGGA 180
Db 37867 ACCTATCAGGCAATATTTGGCATTGACTTTTATCAAAAATATGTACTTCGGAGATGGA 37808

Qy 181 ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGGAACTCTCCGTAGCTCATTTCCC 240
Db 37807 ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGGAACTCTCCGTAGCTCATTTCCC 37748

Qy 241 AGGTACATCCGTGATCTGCTGCAGCTGTAGTATTAGTATACATAACAATGTTAACTCA 300
Db 37747 AGGTACATCCGTGATCTGCTGCAGCTGTAGTATTAGTATACATAACAATGTTAACTCA 37688

Qy 301 TTCAGCAAACTACAAAGTGATGATGATGTGCAAGACAGAGAGGAGTGTATTATC 360
Db 37687 TTCAGCAAACTACAAAGTGATGATGATGTGCAAGACAGAGAGGAGTGTATTATC 37628

Qy 361 ATCAGCTAGTAGGAAATAGAACAGATCTTGTGCAAGAGGCAAGTGTCAAGTTGAGGAG 420
Db 37627 ATCAGCTAGTAGGAAATAGAACAGATCTTGTGCAAGAGGCAAGTGTCAAGTTGAGGAG 37568

Qy 421 GGAGAGGAAAGCCAAAGGGCTGAATGTTTATGAACTAGGGCAAAACTGGA 480
Db 37567 GGAGAGGAAAGCCAAAGGGCTGAATGTTTATGAACTAGGGCAAAAGCTGGA 37508

Qy 481 TACAATGTAAGCAGCTCTTTGAGCTGTAGCAGCAGCTTTGCCGGGAATCGGAAGCACA 540
Db 37507 TACAATGTAAGCAGCTCTTTGAGCTGTAGCAGCAGCTTTGCCGGGAATCGGAAGCACA 37448

Qy 541 CAGGACGGAAGCAGAGAGACATGAGTGACATAAACTGGAAAAAGCCTCAGGAGCAAAACA 600
Db 37447 CAGGACGGAAGCAGAGAGACATGAGTGACATAAACTGGAAAAAGCCTCAGGAGCAAAACA 37388

Qy 601 GTCAGCGAAGGGGGTGTCTCTGCTACCTCCCATGTCAATTTCAACCCCTTCAGAG 660
Db 37387 GTCAGCGAAGGGGGTGTCTCTGCTACCTCCCATGTCAATTTCAACCCCTTCAGAG 37328

Qy 661 CCCCTTACTCTTTCATTTGACTGAGTGAATATGCTTGAACCTTTCCCTTCATTA 720
Db 37327 CCCCTTACTCTTTCATTTGACTGAGTGAATATGCTTGAACCTTTCCCTTCATTA 37268

Qy 721 ATAAAGTTTTGCAATTCATATTCCTGCTGTCTCGTGGAGA 762
Db 37267 ATAAAGTTTTGCAATTCATATTCCTGCTGTCTCGTGGAGA 37226

RESULT 9
AC120839/c 242865 bp DNA linear HTG 05-MAR-2004
LOCUS Pan troglodytes clone rp43-5n22, WORKING DRAFT SEQUENCE, 3 ordered
DEFINITION pieces.
AC120839
VERSION AC120839.38 GI:45120242
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Pan troglodytes (chimpanzee)

Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 242865)
Zhou,L., Fu,Y., Shi,R., Wu,J., Shaul,S., Eichler,E. and Roe,B.A.
Pan troglodytes BAC Clone rp43-5n22
Unpublished
2 (bases 1 to 242865)
Zhou,L., Fu,Y., Shi,R., Wu,J., Shaul,S., Eichler,E. and Roe,B.A.
Direct Submission
Submitted (13-MAY-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 242865)
Zhou,L., Fu,Y., Shi,R., Wu,J., Shaul,S., Eichler,E. and Roe,B.A.
Direct Submission
Submitted (05-MAR-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Mar 5, 2004 this sequence version replaced gi:42415615.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 4245: contig of 4245 bp in length
* 4246 4345: gap of unknown length
* 4346 106118: contig of 101773 bp in length
* 106119 106219: gap of unknown length
* 106219 242865: contig of 136647 bp in length.
FEATURES
source
1. 242865
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="rp43-5n22"
/clone_lib="RPCI - 43 Male Chimpanzee BAC Library"

ORIGIN
Query Match 97.7%; Score 744.4; DB 2; Length 242865;
Best Local Similarity 98.6%; Pred. No. 2.4e-216;
Matches 751; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ATGTCGGCGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 60
Db 128482 ATGTCGGCGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 128423

Qy 61 GAGCAAGCGTTCGAAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGAACAAC 120
Db 128422 GAGCAAGCGTTCGAAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGAACAAC 128363

Qy 121 ACCTATCAGGCAATATTTGGCATTGACTTTTATCAAAAATATGTACTTCGGAGATGGA 180
Db 128362 ACCTATCAGGCAATATTTGGCATTGACTTTTATCAAAAATATGTACTTCGGAGATGGA 128303

Qy 181 ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGGAACTCTCCGTAGCTCATTTCCC 240
Db 128302 ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGGAACTCTCCGTAGCTCATTTCCC 128243

Qy 241 AGGTACATCCGTGATCTGCTGCAGCTGTAGTATTAGTATACATAACAATGTTAACTCA 300
Db 128242 AGGTACATCCGTGATCTGCTGCAGCTGTAGTATTAGTATACATAACAATGTTAACTCA 128183

Qy 301 TTCAGCAAACTACAAAGTGATGATGATGTGCAAGACAGAGAGGAGTGTATTATC 360
```

[illegible]

Db	86875	ATAACGTTTGC	AATTCATCAT	TGCTCCTG	TCTCAT	CGAGA	86916	
RESULT 11	BC003617	3070 bp	mRNA	linear	PRI 29-JUN-2004			
LOCUS	BC003617	3070 bp	mRNA	linear	PRI 29-JUN-2004			
DEFINITION	1.	mRNA (cDNA clone MGC:1654 IMAGE:3506585), complete cds.						
ACCESSION	BC003617							
VERSION	BC003617.2	GI:33870607						
KEYWORDS	MGC.							
SOURCE	Homo sapiens							
ORGANISM	Homo sapiens (human)							
REFERENCE	1	(bases 1 to 3070)						
AUTHORS	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzyzanski, M.I., Skalska, U., Smailus, D.E., Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.							
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences							
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26), 16899-16903 (2002)						
PUBMED	12477932							
REFERENCE	2	(bases 1 to 3070)						
AUTHORS	Strausberg, R.							
TITLE	Direct Submission							
JOURNAL	Submitted (26-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA							
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov							
COMMENT	On Aug 19, 2003 this sequence version replaced gi:13177663. Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/BTP cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amadon@systemsbio.org Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting							
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 7 Row: a Column: 8 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19923230. Location/Qualifiers 1. .3070 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:1654 IMAGE:3506585" /tissue_type="Kidney, renal cell adenocarcinoma" /clone_lib="NIH_MGC_14" /lab_host="DH10B-R" /note="Vector: pOTB7"							
source	1. .3070							
	/gene="RAB6A"							
	/note="synonym: RAB6"							
	/db_xref="LocusID:5870"							
	/db_xref="MIM:179513"							
	418. .1044							
	/gene="RAB6A"							
	/codon_start=1							
	/product="RAB6A, member RAS oncogene family, isoform a"							
	/protein_id="AAH03617.1"							
	/db_xref="GI:13177664"							
	/db_xref="LocusID:5870"							
	/db_xref="MIM:179513"							
	/translation="MSGGDFGNPLRKFVLVFLGEQSVGKTSILTRPMYDFEDNTYQA TIGIDFLSKYMLIEDRTIRLQMDTQAGFRSLIPSYIRDSAAVVMVYDITNNVSQ QTTWKIDDVETERGSDVIIMLVNKTDLADKQVSIIEGKAKELNVMVFIETSAKAG YNVKQLFRVAAALPGMESTQDRSRDMIDIKLEKPOEPVSEGGCSC"							
CDS								
ORIGIN								
	Query Match 89.8%; Score 684; DB 9; Length 3070; Best Local Similarity 93.9%; Pred. No. 4.4e-198; Matches 727; Conservative 0; Mismatches 35; Indels 12; Gaps 1;							
	1 ATGTCGCGCGGCGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTTCCTCGGG 60							
	418 ATGTCACGCGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTTCCTCGGG 477							
	61 GAGCAAAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTGCAAC 120							
	478 GAGCAAAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTGCAAC 537							
	121 ACCTATCAGCAATAATTTGCATGACTTTTATCAAAAATCTATGTAATTTGGAGGATGGA 180							
	538 ACCTATCAGGCAACAATTTGGCATTGACTTTTATCAAAAATCTATGTAATTTGGAGGATCGA 597							
	181 ACAATCGGGCTTCGGCTGTGGGATACGCGGGTCAGGAACGTCCTCGTAGCTCATTTCCC 240							
	598 ACAATCAGGCTTCAGCTGTGGGATACGCGGGTCAGGAACGTTTCGTAGCTCATTTCCC 657							
	241 AGGTACATCCGTGATTCGTCGAGCTGTAGTGTATGATTTACGATATCAAAATGTTAACTCA 300							
	658 AGTTACATCCGTGATTCGTCGAGCTGTAGTGTATGATTTACGATATCAAAATGTTAACTCA 717							
	301 TTCACGCAAACTCAAAAAGTCGATGTGATGTGATGTGATGTGATGTGATGTGATGTATC 360							
	718 TTCACGCAAACTCAAAAAGTCGATGTGATGTGATGTGATGTGATGTGATGTGATGTATC 777							
	361 ATCACGCTAGTAGGAATAAGACAGATCTTGTGTCGACAAAGAGGCAAGTGTGATGTGAGGAG 420							
	778 ATCATGCTAGTAGGAATAAAACAGATCTTGTGCTGACAAAGAGGCAAGTGTCAATTGAGGAG 837							
	421 GGAGAGAGAAAGCCAAAGGCTGAATGTTACGTTTATTTGAAACTAGGCGAAAACCTGGA 480							
	838 GGAGAGAGAAAGCCAAAGAGCTGAATGTTATGTTATTTGAAACTAGTGTGAAAAGCTGGA 897							
	481 TACAATGTTAAAGCAGCTCTTTTCGACGCTGTAGCAGCAGCTTTGCGGGGAATGGAAGCACA 540							
	898 TACAATGTTAAAGCAGCTCTTTTCGACGCTGTAGCAGCAGCTTTGCGGGGAATGGAAGCACA 957							
	541 CAGGACGGAAGCAGAGAAAGACATGAGTGACATAAACTGGAAAAGCCCTCAGGAGCAACA 600							
	958 CAGGACGAGAGCAGAGAAGATATGATTGACATAAACTGGAAAAGCCCTCAGGAGCAACA 1017							
	601 GTCAGCGAAGGGGTTGTTCTCTGCTACTCTCCCATGTCATCTTCAACCCCT----- 650							
	1018 GTCAGTGAAGGAGGCTGTCTCTGCTAATCTCCCATGTCATCTTCAACCCCTTCTTCAGAAGC 1077							
	651 --TCCTCAGAGCCCCCTTACTCTTTCATGATGCTGAGTGTGTAATTTGCGTTGAACCTT 708							
	1078 TCATGCTTTGGCCCCCTTACTCTTTCATGATGCTGAGTGTGTAATTTGCGTTGAACCTT 1137							
	709 TTCCTTTTCAATTAATAACGTTTTCGCAATTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 762							


```
Db 1138 TTCCTTCAGTAATACGTAATTCGAATTCATCATCTGCTGCTGCGTGAGA 1191

RESULT 12
LOCUS BC068486
DEFINITION Homo sapiens RAB6A, member RAS oncogene family, mRNA (cDNA clone
ACCESSION BC068486
VERSION BC068486.1 GI:46249770
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3100)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Faney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalios, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 3100)
Straussberg, R.
Direct Submission
Submitted (02-APR-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAK Plate: 168 Row: 1 Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19923230.
Location/Qualifiers
1..3100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87150 IMAGE:30342562"
/tissue_type="Placenta, normal"
/clone_lib="NIH MGC_147"
/lab_host="DH10B"
/note="Vector: pBluescript"

FEATURES
source
1..3100
/gene="RAB6A"
/note="synonym: RAB6"
/db_xref="LOCUSID:5870"
/db_xref="MIM:179513"
463..1089
/gene="RAB6A"
/codon_start=1
/product="RAB6A protein"
/protein_id="AAH68486.1"
/db_xref="GI:46249771"
/db_xref="LOCUSID:5870"
/db_xref="MIM:179513"
/misc_feature
/gene="RAB6A"
/note="RAB6; Region: Rab subfamily of small GTPases"
/db_xref="CDD:cd00154"

ORIGIN
Query Match 89.8%; Score 684; DB 9; Length 3100;
Best Local Similarity 93.9%; Pred. NO. 4.4e-198;
Matches 727; Conservative 0; Mismatches 35; Indels 12; Gaps 1;
QY 1 ATGTCCGCGCGGAGACTTCGGGAATCCGCTGAGGAATTCAGCTGGTGTCTCTGGGG 60
DB 463 ATGTCCAGCGCGGAGACTTCGGGAATCCGCTGAGGAATTCAGCTGGTGTCTCTGGGG 522
QY 61 GAGCAAAACGCTTGCAAGACATCTTTGATCACCAGATTCAGTATGACAGATTTTCACAAC 120
DB 523 GAGCAAAACGCTTGCAAGACATCTTTGATCACCAGATTCAGTATGACAGATTTTCACAAC 582
QY 121 ACCTATCAGGCAATAATTGGCATTGACATTTTATCAAAAACATGATGATTTGGAGATGA 180
DB 583 ACCTATCAGGCAATAATTGGCATTGACATTTTATCAAAAACATGATGATTTGGAGATGA 642
QY 181 ACAATCGGCTTCGGCTGCGGATACGCGGCTGAGGAACGCTCCGCTAGCTCATTTCCC 240
DB 643 ACAATCGGCTTCAGCTGCGGATACGCGGCTGAGGAACGCTTCGCTAGCTCATTTCCC 702
QY 241 AGGTACATCCGCTGATTCGCTGAGCTGATGATTTACGATATCACAAATGTTAACTCA 300
DB 703 AGGTACATCCGCTGATTCGCTGAGCTGATGATTTACGATATCACAAATGTTAACTCA 762
QY 301 TTCAGCAAACTACAAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 763 TTCAGCAAACTACAAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 822
QY 361 ATCAGCTAGTAGGAATAGAACAGATCTTCTGACAGAGGCAAGTCTCAGTTGAGGAG 420
DB 823 ATCATGCTAGTAGGAATAGAACAGATCTTCTGACAGAGGCAAGTCTCAGTTGAGGAG 882
QY 421 CGAGAGGAAAGCCAAAGGCTGAATGTTATGATTTTATGAAACTAGGCAAAACTGGA 480
DB 883 GGAGAGGAAAGCCAAAGGCTGAATGTTATGATTTTATGAAACTAGTGCAGAAAGCTGGA 942
QY 481 TACAATGTAAAGCAGCTCTTTTCGACGCTGATGATGATGATGATGATGATGATGATGATGAT 540
DB 943 TACAATGTAAAGCAGCTCTTTTCGACGCTGATGATGATGATGATGATGATGATGATGATGAT 1002
QY 541 CAGAGCGAAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 1003 CAGAGCGAAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1062
QY 601 GTCAGCAAGGCGGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
DB 1063 GTCAGTGAAGAGAGCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122
QY 651 --TCCTCAGAAAGCCCTTACTCTTTCAATGATGATGATGATGATGATGATGATGATGATGAT 708
```

Db 1123 TCACGTCTTTGGCCCCCTTACTCTTTTCATTGACTGAGTGTAATTTGGCTTGAACCTT 1182

Qy 709 TTCCTTCATTAAATTAACGTTTGGCAATTCATTCATTCGCTGCTGCTGCGTGGAGA 762

Db 1183 TTCCTTCAGTAATAACGATTATGCAATTCATTCATTCGCTGCTGCTGCGTGGAGA 1236

RESULT 13

AK057157

LOCUS

DEFINITION

Homo sapiens CDNA FLJ32595 fis, clone SPLEN2000505, highly similar

to RAS-RELATED PROTEIN RAB-6.

ACCESSION

AK057157.1 GI:16552747

VERSION

oligo capping; fis (full insert sequence).

KEYWORDS

Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,

Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,

Sekine, M., Ohayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,

Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,

Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,

Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,

Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,

Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,

Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,

Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,

Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,

Yosida, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A.,

Hara, H., Tanase, T., Nomura, Y., Togliani, S., Komai, F., Hara, R.,

Takeuchi, K., Arita, M., Imase, N., Musashino, K., Yuuki, H., Oshima, A.,

Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,

Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,

Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,

Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,

Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,

Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,

Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,

Ozaki, K., Hiraoka, N., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,

Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,

Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K.,

Nakajima, T., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,

Oyama, H., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,

Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,

Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,

Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.

Complete sequencing and characterization of 21,243 full-length

human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

14702039

2

Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,

Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,

Isono, F., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,

Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,

Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,

Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K.,

Masuho, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

Unpublished

3 (bases 1 to 3120)

Isogai, T., Otsuki, T. and Sugiyama, T.

Direct Submission

Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,

Genomics Laboratory; 1532-3 Yano Kisarazu, Chiba 292-0812, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

FEATURES

source

Location/Qualifiers

1..3120

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="SPLEN2000505"

/tissue_type="spleen"

/clone_lib="SPLEN2"

/note="Cloning vector: pME18SFL3"

487..1113

/note="unnamed protein product"

/codon_start=1

/protein_id="BAB71371.1"

/db_xref="GI:16552748"

/translation="MSTGGDPGNPLRKFVLFLGEQSVGKTSLTRWYDSFDNTYQA

TIGIDFLSKTWYLEDRTIRLQWDTAGQFRSLIPSYIRDSAAAANVVDITWNVSFO

QTKWIDDVTERGSDVIMLVGNKTDLADKRQVSIIEGGRKAKELNVMFIETSAKAG

YNVKQLFRVAAALPGMESTQDRSDIMIDIKLEKPEQPVSEGGCSC"

ORIGIN

Query Match 89.8%; Score 684; DB 9; Length 3120;

Best Local Similarity 93.9%; Pred. No. 4.4e-198; Indels 12; Gaps 1;

Matches 72; Conservative 0; Mismatches 35;

Qy 1 ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGAGCTGGTGTTCCTGGGG 60

Db 487 ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGAGCTGGTGTTCCTGGGG 546

Qy 61 GAGCAAAAGCGTTGCAAAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGCACAAC 120

Db 547 GAGCAAAAGCGTTGCAAAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGCACAAC 606

Qy 121 ACCTATCAGCAATAATTTGCATTGACCTTTTATCAAAAATATGATCTTGGAGGATGGA 180

Db 607 ACCTATCAGCAATAATTTGCATTGACCTTTTATCAAAAATATGATCTTGGAGGATGGA 666

Qy 181 ACAATCGGGCTTCGGCTGTGGGATACGCGGGTCAGGAACGTCCTCCGTAGCCTCATTTCCC 240

Db 667 ACAATCAGGCTTCAGCTGTGGGATACGCGGGTCAGGAACGTCCTCCGTAGCCTCATTTCCC 726

Qy 241 AGGTACATCGTGTATCTGCTGAGCTGTAGTAGTTTACAGATATCAGAAATGTTAACTCA 300

Db 727 AGTTACATCGTGTATCTGCTGAGCTGTAGTAGTTTACAGATATCAGAAATGTTAACTCA 786

Qy 301 TTCCAGCAAACTACAAAGTCGATTGATGTGTAGATGATGATGATGATGATGATGATGATGAT 360

Db 787 TTCCAGCAAACTACAAAGTCGATTGATGTGTAGATGATGATGATGATGATGATGATGATGAT 846

Qy 361 ATCAGCTAGTAGGAAATAGAACAGATCTTCTGCTGACAAAGGCGCAAGTGTCTAGTTGAGGAG 420

Db 847 ATCATGCTAGTAGGAAATAGAACAGATCTTCTGCTGACAAAGGCGCAAGTGTCTAGTTGAGGAG 906

Qy 421 GGAGAGAGGAAAGCAAGGCTGAATGTTACGTTTATGAACTAGGCGCAAACTGGA 480

Db 907 GGAGAGAGGAAAGCAAGGCTGAATGTTATGTTTATGAACTAGTGTGCAAAAGCTGGA 966

Qy 481 TACAATCTAAAGCAGCTCTTTTCGACGTTGAGCAGCAGCTTTCCGCGGAATGGAAGACACA 540

Db 967 TACAATCTAAAGCAGCTCTTTTCGACGTTGAGCAGCAGCTTTCCGCGGAATGGAAGACACA 1026

Qy 541 CAGGACGGAAGCAGAGAGACATGAGTGACATATAAACTGGAAGGCTTCAGGAGCAAAACA 600

Db 1027 CAGGACGGAAGCAGAGAGACATGAGTGACATATAAACTGGAAGGCTTCAGGAGCAAAACA 1086

Qy 601 GTACGAGAGGCGGTTGTTCTCTGCTACTCTCCCATGTCTCAACCTT----- 650

Db 1087 GTACGAGAGGCGGTTGTTCTCTGCTACTCTCCCATGTCTCAACCTTCTTCAGAGC 1146

Qy 651 -TCCTCAGAAAGCCCTTACTCTTTTCATTGACTGCGAGTGTGAATATTTGGCTTGAACCTT 708

Db	1094	GTCAAGTGAAGGAGGCTGTTCCTCGCTAATTCCTCATGTCTCATCTTCAACCTTTCTTCGAAGC	1155
Qy	651	--TCCTCAGAAGCCCTTACTCTTTTCATTTGACTGCAGTGTGAATATTTGGCTTTGAACCTT	708
Db	1154	TCAGTCTGTTGGCCCTTACTCTTTTCATTTGACTGCAGTGTGAATATTTGGCTTTGAACCTT	1213
Qy	709	TTCCCTTCATTAATAACGTTTGGCAATTCATCATTTGCTGCTGCTCTCGTGGAGA	762
Db	1214	TTCCCTTCAGTAATACGATTGCAATTCATCATTTGCTGCTGCTCTCGTGGAGA	1267
RESULT 15			
BC019118			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 39 Row: a Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.

FEATURES

Location/Qualifiers

1..1770
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:29434 IMAGE:3711087"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

gene

1..1770
/gene="Rab6"
/db_xref="LocusID:19346"
/db_xref="MGI:894313"
271..897

CDS

/gene="Rab6"
/codon_start=1
/product="Rab6 protein"
/protein_id="AAH19118.1"
/db_xref="GI:17512290"
/db_xref="LocusID:19346"
/db_xref="MGI:894313"

/translation="MSGAGDFGNLRFKLVFLGEQSVGKTSLLTRWYDSFDNTYOA
TIGIDFLSKWYLEDRTIQLWDTAGQERFLPSIYRDSAAVYVDITNVNSFO
QTTWIDVTRGSDVIIMLVGNKTLADKROVSIIEGKAKELNMFIEISAKAG
YNVQLFRRVAALPGMESTQDRSDMDIDIKLEKPOEQPVNEGSCS"
310..792

misc_feature

/gene="Rab6"
/note="RAB; Region: Rab subfamily of small GTPases"
/db_xref="CDD:smart00175"

ORIGIN

Query Match 83.8%; Score 638.2; DB 10; Length 1770;
Best Local Similarity 90.3%; Pred No. 4.6e-184;
Matches 698; Conservative 0; Mismatches 63; Indels 12; Gaps 1;

Qy 1 ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTCTTCCTGGGG 60
Db |
271 ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTCTTCCTGGGA 330
Qy |
61 GAGCAAGCGTTGCAAGACATCTTGATCACCAGATTGAGTATGACAGTTTTCACAAAC 120
Db |
331 GAGCAGACGTTGGAAGACGCTCTTGATCACCAGATTGATGATGACAGTTTTCACAAAC 390
Qy |
121 ACCTATCAGGCAATATTTGGCATTGACATTTTATCAAAATATGATCTGAGGATGGA 180
Db |
391 ACCTATCAGGCAACAAATTTGGCATTGACATTTATCAAAACAAATGATCTTGAGGATAGA 450
Qy |
181 ACAATCGGGCTTCGGCTCTGGGATACGGCGGTTCAGGAACGCTCCGCTAGCCTCATTTCCC 240
Db |
451 ACCATCAGGCTGCAGCTGCGGATACGCGGTTCAGGAACGCTTCGCTAGCCTCATTTCCC 510
Qy |
241 AGGTACATCCGTGATCTGCTGACGCTGTAGTAGTTACGATATCAAAATGTTAACTCA 300
Db |
511 AGTTACATCCGTGACTCTGCTGCAGCCGTAGTAGTTTACGATATCACAAATGTTAACTCA 570
Qy |
301 TTCCAGCAAACTACAAAGTGGATTGATGTCAGAACAGAGAGGAAGTGTGTTATC 360
Db |
571 TTCCAGCAAACTACAAATGGAATGATGATGTCAGAACAGAGAGGAAGTGTGTTATC 630
Qy |
361 ATCAGCTAGTAGGAAATAGAACAGATCTTGCTGCAAGAGGCAAGTGTGAGTTGAGGAG 420
Db |
631 ATCATGCTAGTAGGAAATATAAACAGATCTTGCTGATAGAGGCAAGTGTCAATTTGAGGAG 690

Qy 421 GGAGAGAGGAAAGCCAAAGGCTGAATGTTTACGTTTATTGAAACTAGGCAAAACTGGA 480
Db |
691 GGAGAGAGGAAAGCCAAAGAGCTGAATGTTTATTGAAACCACTGCAAAAGCAGGA 750
Qy 481 TACAATGTAAGCAGCTCTTTTCGACGCTGTAGCAGCAGCTTTGCGGGGAATGGAAAGCACA 540
Db |
751 TACAATGTAAGCAGCTCTTTTCGCGCGTGTGCTGCAGCTTTACCTGGAATGGAAAGCACA 810
Qy 541 CAGGACGGAAGCAGAGACATGATGACATATAAACTGAAAAGCCTCAGGAGCAACA 600
Db |
811 CAGGACGGAAGCAGAGACATGATGACATATAAACTGAAAAGCCTCAGGAGCAACA 870
Qy 601 GTCAATGTAAGCAGCTCTTTTCGCTGCTACTCTCCATGTCATCTTCAACCTTCTCTCAGAAG 660
Db |
871 GTCAATGTAAGCAGCTCTTTTCGCTGCTACTCTCCATGTCATCTTCAACCTTCTCTCAGAAG 930
Qy 661 C-----CCCCCTTACTCTTTCAATGACTGCACTGTGAATATTGGCTTGAACCTT 708
Db |
931 CTCACCTGCTTTGGCCCCCATACTCTTTCAATGACTGCACTGTGAATATTGGCTTGAACCTT 990
Qy 709 TTCCCTTCATTAATAACGTTTTCGAAATTCATCATTCGCTGCTGCTGCTGAG 761
Db |
991 TTCCCTTCAGTAATAACGTTTTCGAAATTCATCATTCGCTGCTGCTGCTGAG 1043

Search completed: April 25, 2005, 10:41:08

Job time : 3402 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 05:41:05 ; Search time 475 Seconds
(without alignments)
9496.511 Million cell updates/sec

Title: US-09-441-857-11
Perfect score: 762
Sequence: 1 atgtccggcgaggagactt.....tgcgtcgtctcgtggaga 762

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	762	100.0	762	3	Aaz93839
2	761	99.9	3084	5	ABX71303
C 3	759.4	99.7	349981	10	ADC87619
C 4	759.4	99.7	349989	10	ADC86916
5	684	89.8	1460	3	Aaz93840
6	684	89.8	3079	12	ADN03693
7	672	88.2	719	3	Aaz93834
8	639.2	83.9	3745	12	ADL12658
9	637.8	83.7	3118	3	AAFI5989
10	609.2	79.9	739	3	Aaz93836
11	585.6	76.9	624	3	Aaz93841
12	559.6	73.4	718	3	Aaz93835
13	559.6	73.4	740	3	Aaz93837
14	526	69.0	573	3	Aaz93838
15	489.8	64.3	3195	5	AA87692
16	401	52.6	2456	10	ADA52862
17	384.8	50.5	1219	4	AAK52711
18	384.8	50.5	1461	4	AAK51727
19	383.2	50.3	1266	12	ADP07320
20	383.2	50.3	1266	13	ADR25795

21	383.2	50.3	1589	2	AAZ11736
22	382	50.1	443	8	ABX42190
23	352.2	46.2	3826	10	ADC86930
24	341.8	44.9	424	9	ACH18328
25	296.4	38.9	1188	4	ABL11393
C 26	296.4	38.9	3188	4	ABL11392
C 27	286.8	37.6	477	5	ABA20123
28	276.8	36.3	642	5	AA87691
29	268.4	35.2	1074	3	AAC33337
30	251.6	33.0	523	13	ACN59588
31	250	32.8	636	10	ADC76216
32	250	32.8	636	10	ADK55884
33	247	32.4	1355	3	AAAC46575
34	246.2	32.3	624	3	AAAC42719
35	245.2	32.2	928	6	ABI99630
36	244.4	32.1	926	3	AAFI14485
37	236	31.0	495	3	AAAC1466
38	227	29.8	646	10	ADD16517
39	227	29.8	646	11	ADM44783
40	227	29.8	647	10	ADC75668
41	227	29.8	723	10	ADC76194
42	227	29.8	723	10	ADK58490
43	216	28.3	574	13	ADR65698
44	213	28.0	645	3	AAAC3018
45	211.8	27.8	791	13	ADR65699

ALIGNMENTS

RESULT 1
AAZ93839
ID AAZ93839 standard; DNA; 762 BP.
XX
AC AAZ93839;
XX 29-AUG-2000 (first entry)
XX WTH3 coding sequence.
XX WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;
KW antibody; immunogen; mutation; detection; therapy; human; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 1..762
FT /*tag= a
FT FT /product= "WTH3"
XX WO200029625-A1.
XX PD 25-MAY-2000.
XX 18-NOV-1999; 99WO-US027630.
XX PR 18-NOV-1998; 98US-0108994P.
PA (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.
PI Duffy HX, Shan J, Yuan L, Budman D, Calabro A;
DR WPI; 2000-387828/33.
XX P-PSDB; AAY83649.
PT Novel nucleic acids encoding wth3 protein useful for inhibiting multiple
PT drug resistance in tumor cells for identifying the modulators of drug
XX resistance.
XX Claim 3; Page 62-64; 82pp; English.
XX Nucleic acids encoding WTH3 and RAB6 were isolated by Methyl- (or mutant)
CC Differential Display (MDD) methods and can be used in methods for

Db 696 AGGTACATCCGATTTCTCTGTCAGCTGTAGTATTACGATATACAAATGTTAACTCA 755
QY 301 TTCCAGCAAACTACAAAGTGGATTGATGATGTGAGAAACAGAAAGAGGAGTGTATTC 360
Db 756 TTCCAGCAAACTACAAAGTGGATTGATGATGTGAGAAACAGAAAGAGGAGTGTATTC 815
QY 361 ATCAGCTAGTAGGAAATAGAACAGATCTTGTCTGACAGAGGCAAGTGTCTGAGGAG 420
Db 816 ATCAGCTAGTAGGAAATAGAACAGATCTTGTCTGACAGAGGCAAGTGTCTGAGGAG 875
QY 421 GGAGAGAGAAAGCCAAAGGGCTGAATGTTAGCTTTATTGAACTAGGCAAAACTGGA 480
Db 876 GGAGAGAGAAAGCCAAAGGGCTGAATGTTAGCTTTATTGAACTAGGCAAAACTGGA 935
QY 481 TACAATGTAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGAAAGCACA 540
Db 936 TACAATGTAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGAAAGCACA 995
QY 541 CAGGACGGAAGCAGAGAACATGAGTGACATATAAACTGGAAGAGCCTCAGGAGCAACA 600
Db 996 CAGGACGGAAGCAGAGAACATGAGTGACATATAAACTGGAAGAGCCTCAGGAGCAACA 1055
QY 601 GTACGCGAAGGGGTTGTTCTGCTACTCTCCATGTCATCTCAACCTTCTCAGAG 660
Db 1056 GTACGCGAAGGGGTTGTTCTGCTACTCTCCATGTCATCTCAACCTTCTCAGAG 1115
QY 661 CCCCTTACTCTTTTCAATGACGTGCAATGTAATATTGGCTTGAACCTTTTCCTTCATTA 720
Db 1116 CCCCTTACTCTTTTCAATGACGTGCAATGTAATATTGGCTTGAACCTTTTCCTTCATTA 1175
QY 721 ATAAGCTTTTGAATTCATATGCTGCTGTCTGCTGTGAG 761
Db 1176 ATAAGCTTTTGAATTCATATGCTGCTGTCTGCTGTGAG 1216

RESULT 3
ADC87619/c
ID ADC87619 standard; DNA; 349981 BP.
XX
AC ADC87619;
XX
DT 01-JAN-2004 (first entry)
DE Human GPCR related polynucleotide SEQ ID NO:2072.
XX
KW ds; human; GPCR; guanosine triphosphate-binding protein coupled receptor;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN EP1270724-A2.
XX
PD 02-JAN-2003.
XX
PF 18-JUN-2002; 2002BP-00013517.
XX
PR 18-JUN-2001; 2001JP-00246789.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX WPI; 2003-315783/31.
XX
PT New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.
XX
PS Disclosure; SEQ ID NO 2072; 28pp; English.
XX
CC The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of

CC the invention may have a use in gene therapy. The polynucleotide and
CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The protein
CC sequences shown in ADC87618-ADC87623 represent polynucleotide sequences
CC related to the invention.
XX
SQ Sequence 349981 BP; 100900 A; 78560 C; 77396 G; 92823 T; 0 U; 302 Other;
Query Match 99.7%; Score 759.4; DB 10; Length 349981;
Best Local Similarity 99.9%; Pred. No. 7.3e-225;
Matches 760; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTCGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTGTCTCTGGGG 60
Db 45699 ATGTCGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTGTCTCTGGGG 45640
QY 61 GAGCAACGCTTGCNAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGCACAC 120
Db 45639 GAGCAACGCTTGCNAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGCACAC 45580
QY 121 ACCTATCAGGCAATAATTGGCAATTGACTTTTATCAAAACTATGTACTTGGAGATGGA 180
Db 45579 ACCTATCAGGCAATAATTGGCAATTGACTTTTATCAAAACTATGTACTTGGAGATGGA 45520
QY 181 ACAATCGGGCTTCGGCTGTGGGATACGCGGGTCAGGAACGCTCTCCGTAGCCTCAATCCC 240
Db 45519 ACAATCGGGCTTCGGCTGTGGGATACGCGGGTCAGGAACGCTCTCCGTAGCCTCAATCCC 45460
QY 241 AGGTACATCCGTGATTCTGCTGAGCTGTAGTATTAGGATATACAAATGTTAACTCA 300
Db 45459 AGGTACATCCGTGATTCTGCTGAGCTGTAGTATTAGGATATACAAATGTTAACTCA 45400
QY 301 TTCCAGCAAACTACAAAGTGGATGATGTGACAGACAGAAAGAGGAGTGTATTC 360
Db 45399 TTCCAGCAAACTACAAAGTGGATGATGTGACAGACAGAAAGAGGAGTGTATTC 45340
QY 361 ATCAGCTAGTAGGAAATAGAACAGATCTTCTGCAAGAGGCAAGTGTCTAGTTGAGGAG 420
Db 45339 ATCAGCTAGTAGGAAATAGAACAGATCTTCTGCAAGAGGCAAGTGTCTAGTTGAGGAG 45280
QY 421 GGAGAGGAAAGCCAAAGGGCTGAAATGTTAGCTTTATTGAACTTAGGGCAAAAGCTGA 480
Db 45279 GGAGAGGAAAGCCAAAGGGCTGAAATGTTAGCTTTATTGAACTTAGGGCAAAAGCTGA 45220
QY 481 TACAATGTAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTTCCGGGAATGGAAGCACA 540
Db 45219 TACAATGTAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTTCCGGGAATGGAAGCACA 45160
QY 541 CAGGACGGAAGCAGAGAACATGAGTGACATAAACTGGAAGAGCCTCAGGAGCAAAACA 600
Db 45159 CAGGACGGAAGCAGAGAACATGAGTGACATAAACTGGAAGAGCCTCAGGAGCAAAACA 45100
QY 601 GTCAGGAAAGGGGTTGTTCTGCTACTCTCCATGTCATCTTCAACCTTCTCAGAG 660
Db 45099 GTCAGGAAAGGGGTTGTTCTGCTACTCTCCATGTCATCTTCAACCTTCTCAGAG 45040
QY 661 CCCCTTACTCTTTTCAATGACGTGCAATGTAATATTGGCTTGAACCTTTTCCCTTCATTA 720
Db 45039 CCCCTTACTCTTTTCAATGACGTGCAATGTAATATTGGCTTGAACCTTTTCCCTTCATTA 44980
QY 721 ATAAGCTTTTGAATTCATATGCTGCTGTCTGCTGTGAG 761
Db 44979 ATAAGCTTTTGAATTCATATGCTGCTGTCTGCTGTGAG 44939

RESULT 4
ADC86916/c
ID ADC86916 standard; DNA; 349989 BP.
XX
AC ADC86916;
XX
DT 01-JAN-2004 (first entry)

CC drug resistance locus can be altered and give rise to altered expression
CC patterns of that multiple drug resistance locus. Nucleic acids
CC corresponding to the identified loci such as WTH3 and RAB6 nucleic acids
CC can be used as probes for detecting mutations and methylation patterns of
CC those loci. The nucleic acids and their homologues are useful for
CC inhibition of multiple drug resistance and for treating tumors exhibiting
CC multiple drug resistance. They are also useful for detecting and
CC measuring the expression of mRNA from identified genes and for
CC determining suitable therapeutic treatment. Antibodies directed against
CC immunogenic fragments of WTH3 and RAB6 are useful for detecting specific
CC proteins and polypeptides in tissues or body fluids of patients
XX
SQ Sequence 1460 BP; 365 A; 356 C; 358 G; 381 T; 0 U; 0 Other;

Query Match 89.8%; Score 684; DB 3; Length 1460;
Best Local Similarity 93.9%; Pred. No. 1.6e-202;
Matches 727; Conservative 0; Mismatches 35; Indels 12; Gaps 1;

QY 1 ATGTCGGCGGGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTGTCTCTGGGG 60
Db 420 ATGTCACGCGGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTGTCTCTGGGG 479

QY 61 GAGCAAGCGTTGCAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGCACAC 120
Db 480 GAGCAAGCGTTGCAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGCACAC 539

QY 121 ACTATCAGGCAATTAATCGCAATTCGCTTTTATCAAAACATGATGATGAGGATGGA 180
Db 540 ACTATCAGGCAATTAATCGCAATTCGCTTTTATCAAAACATGATGATGAGGATGGA 599

QY 181 ACAATCGGCTTCGGCTGGGATAGCGGGTTCAGGAACGCTCCGTAGCCCTCATPTCCC 240
Db 600 ACAATCAGGCTTCGGCTGGGATAGCGGGTTCAGGAACGCTTCGTAGCCCTCATPTCCC 659

QY 241 AGGTACATCCGTGATTTCTCTGAGCTGTAGTATGATGATGATGATGATGATGATGAT 300
Db 660 AGGTACATCCGTGATTTCTCTGAGCTGTAGTATGATGATGATGATGATGATGATGAT 719

QY 301 TTCCAGCAAACTACAAAGTGGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 720 TTCCAGCAAACTACAAAGTGGATGATGATGATGATGATGATGATGATGATGATGATG 779

QY 361 ATCAGCTAGTAGGAAATAGAACAGATCTTGTCTGACAGAGGCAAGTGTGAGGAG 420
Db 780 ATCATGCTAGTAGGAAATAGAACAGATCTTGTCTGACAGAGGCAAGTGTGAGGAG 839

QY 421 GGAGAGAGAAAGCCAAAGGCTGAAATGTTACGTTTATTTGAACTAGGGCAAAACCTGGA 480
Db 840 GGAGAGAGAAAGCCAAAGGCTGAAATGTTATGTTTATTTGAACTAGTGTCAAAAGCTGA 899

QY 481 TACAAATGTAAGCAGCTCTTTTCGAGCTGTAGCAGCAGCTTTCCCGGGAATGAAAGCACA 540
Db 900 TACAATGTAAGCAGCTCTTTTCGAGCTGTAGCAGCAGCTTTCCCGGGAATGAAAGCACA 959

QY 541 CAGGACGGAAGCAGAGAACATGAGTGACATATAAACTGGAAGGCTCAGGAGCAAAACA 600
Db 960 CAGGACGGAAGCAGAGAACATGAGTGACATATAAACTGGAAGGCTCAGGAGCAAAACA 1019

QY 601 GTCAGGAAAGGGGGTGTTCCTGCTACTCTCCATGTCTCATCTTCAACCCCT----- 650
Db 1020 GTCAGTGAAGAGGGTGTTCCTGCTAAATCTCCCATGTCTCATCTTCAACCCCTTCTCAGAAGC 1079

QY 651 --TCCTCAGAAAGCCCTTACTCTTTCATTTGATGCTGAGTGTGATGATGATGATGATG 708
Db 1080 TCACGTCTTTGGCCCCCTTACTCTTTCATTTGATGCTGAGTGTGATGATGATGATGATG 1139

QY 709 TTCCCTTCAATTAACGCTTTTGGCAATTCATGCTCCCTGTCTCGTGGAGA 762
Db 1140 TTCCCTTCAATTAACGCTTTTGGCAATTCATGCTCCCTGTCTCGTGGAGA 1193

ID ADN03693 standard; cDNA; 3079 BP.
XX
AC ADN03693;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #44.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305105/28.
XX
PT P-PSDB; ADN03694.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 1; SEQ ID NO 87; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 3079 BP; 901 A; 639 C; 636 G; 903 T; 0 U; 0 Other;

Query Match 89.8%; Score 684; DB 12; Length 3079;
Best Local Similarity 93.9%; Pred. No. 2.3e-202;
Matches 727; Conservative 0; Mismatches 35; Indels 12; Gaps 1;

QY 1 ATGTCGGCGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 60
Db 427 ATGTCACGCGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 486

QY 61 GAGCAAGCGTTGCAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGCACAC 120
Db 487 GAGCAAGCGTTGCAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGCACAC 546

QY 121 ACCTATCAGGCAATTAATCGCAATTCGCTTTTATCAAAACATGATGATGAGGATGGA 180
Db 547 ACCTATCAGGCAATTAATCGCAATTCGCTTTTATCAAAACATGATGATGAGGATGGA 606

QY 181 ACAATCGGCTTCGGCTGTGGATGACGGGGTTCAGGAACGCTCCGTAGCCCTCATPTCCC 240
Db 607 ACAATCAGGCTTCAGCTGTGGATGACGGGGTTCAGGAACGCTCCGTAGCCCTCATPTCCC 666

QY 241 AGGTACATCCGTGATTTCTGCTGCACTGTAGTATGATGATGATGATGATGATGATGAT 300
Db 667 AGGTACATCCGTGATTTCTGCTGCACTGTAGTATGATGATGATGATGATGATGATGAT 726

QY 301 TTCCAGCAAACTACAAAGTGGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 727 TTCCAGCAAACTACAAAGTGGATGATGATGATGATGATGATGATGATGATGATGATG 786

QY 361 ATCAGCTAGTAGGAAATAGAACAGATCTTCTGCAAGAGGCAAGTGTGAGTGGAG 420
Db 787 ATCATGCTAGTAGGAAATAGAACAGATCTTCTGCAAGAGGCAAGTGTGAGTGGAG 846

XX 06-MAY-2004 (first entry)
XX Human steroid-induced C3A liver cell cDNA #387.
XX ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX
XX Homo sapiens.
XX US6673549-B1.
XX
XX 06-JAN-2004.
XX 12-OCT-2001; 2001US-00976594.
XX
XX 12-OCT-2000; 2000US-0240409P.
XX
XX (INCY-) INCYTE CORP.
XX Furness LM, Buchbinder JL;
XX WPI; 2004-068610/07.
XX
XX Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
XX Claim 1; SEQ ID NO 387; 141pp; English.
XX
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 3745 BP; 1066 A; 755 C; 745 G; 1179 T; 0 U; 0 Other;

Query Match 83.9%; Score 639.2; DB 12; Length 3745;
Best Local Similarity 90.3%; Pred. No. 2.5e-188;
Matches 699; Conservative 0; Mismatches 63; Indels 12; Gaps 1;
1 ATGTCGGGGGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTTCCTGGG 60
494 ATGTCACGGGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTTCCTGGG 553
61 GAGCAAGCGTTGCAAGACATCTTTGATCACAGATTCCAGTATGACAGATTTTGACAAC 120
554 GAGCAAGCGTTGGAAGACATCTTTGATCACAGATTTCATGATGACAGATTTTGACAAC 613
121 ACCTATCAGGCAATTAATGGCAATTCGATTTTATCAAAACATGATGATGAGGATGGA 180
614 ACCTATCAGGCAATTAATGGCAATTCGATTTTATCAAAACATGATGATGAGGATGGA 673
181 ACAATCGGCTTCGGCTGGGATACGGGGGTGAGGAACGTCTCGTAGCCCTCATTTCC 240
674 ACAGTACGATTCGAATTTATGGGACACAGCAGGTCAAGAGCGGTTTCAGGAGCTTGATTC 733

QY 241 AGGTACATCCGTGATTTCTGCTGCAGCTGTAGTAGTTTACGATATCAGAAATGTTAACTCA 300
DB |||||
734 AGCTACATTCGTGACTCCACTGTGCAGTGTGTTTATGATATCAGAAATGTTAACTCA 793
QY 301 TTCAGCAAACTACAAAGTGAATGATGTAGACAGAGAGAGAGAGAGAGAGAGAGAG 360
DB |||||
794 TTCAGCAAACTACAAAGTGAATGATGTAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 853
QY 361 ATCAGCTAGTAGGAATAGAAACAGATCTTCTGACAGAGAGAGAGAGAGAGAGAGAGAG 420
DB |||||
854 ATCATGTAGTAGGAATAGAAACAGATCTTCTGACAGAGAGAGAGAGAGAGAGAGAGAG 913
QY 421 CGAG 480
DB |||||
914 CGAG 973
QY 481 TACAATGTAAAGCAGCTCTTTCGACGCTGTAGCAGAGAGCTTTGCGGGAGAGAGAGAGAG 540
DB |||||
974 TACAATGTAAAGCAGCTCTTTCGACGCTGTAGCAGAGAGCTTTGCGGGAGAGAGAGAG 1033
QY 541 CAGGAG 600
DB |||||
1034 CAGGAG 1093
QY 601 GTACAG 650
DB |||||
1094 GTACAG 1153
QY 651 --TCTCTCAG 708
DB |||||
1154 TCACTGCTTTGGCGCCCTTACTCTTCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1213
QY 709 TTCCCTTCATTAATAAGCTTTTGGCAATTCATCATTCGCTGCTGCTGCTGAGAGAGAG 762
DB |||||
1214 TTCCCTTCAGTAATAAGCTTTTGGCAATTCATCATTCGCTGCTGCTGCTGAGAGAGAG 1267

RESULT 9
AAFI5989
ID AAFI5989 standard; cDNA; 3118 BP.
AC AAFI5989;
DT 13-MAR-2001 (first entry)
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:424.
XX Human prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antineoplastic; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
XX Homo sapiens.
XX WO2000055174-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005988.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-587513/55.
DR P-PSDB; AAB56786.
XX
XX Prostate cancer associated gene sequences, referred to as prostate cancer

PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.

XX Claim 1; Page 906; 2338pp; English.

XX AAF1566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cyrostatic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, antinefactive, gynaeological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
XX invention

XX Sequence 3118 BP; 938 A; 636 C; 638 G; 902 T; 0 U; 4 Other;

Query Match 83.7%; Score 637.8; DB 3; Length 3118;
Best Local Similarity 90.1%; Pred. No. 6.3e-188;
Matches 697; Conservative 1; Mismatches 64; Indels 12; Gaps 1;

QY 1 ATGTCGGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 60
DB |||||
QY 61 GAGCAAGCGTTGCAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTGTGCAAC 120
DB |||||
QY 475 GAGCAAGCKNTGGAAGACATCTTGTATCACCAGATTCATGTATGACAGTTTGTGCAAC 534
DB |||||
QY 121 ACCATCAGGCAATTAATGGCATTGACATTTTATCAAAAATATGTATCTGGAGATGGA 180
DB |||||
QY 535 ACCATCAGGCAACATTTGGCATTGACATTTTATCAAAAATATGTATCTGGAGATGGA 594
DB |||||
QY 181 ACAATCGGGCTTCGGCTGTGGGATACGGCGGTGAGGAACGTCTCCGTAGCCTCATTCCT 240
DB |||||
QY 595 ACAGTACGATTGCAATTAATGGGACACAGCAGGTCAAGAGCGGTTGAGGAGCTTGTATCT 654
DB |||||
QY 241 AGGTACATCCGTGATTCGCTGCAGCTGTAGTTTACGATATACAAAATGTTAACTCA 300
DB |||||
QY 655 AGCTACATTCGTGACTCCACTGTGGCAGTGTGTATGATATACAAAATGTTAACTCA 714
DB |||||
QY 301 TTCAGCAAACTACAAAGTGATGATGATCTCAGAACAGAGAGGAAGTGTATCTATC 360
DB |||||
QY 715 TTCCAGCAAACTACAAAGTGATGATGATCTCAGAACAGAGAGGAAGTGTATCTATC 774
DB |||||
QY 361 ATCAGCTAGTAGGAATAGAACAGATTTCTGCTGACAAAGGCAAGTGTCAAGTTGAGGAG 420
DB |||||
QY 775 ATCATGCTAGTAGGAATAAAACAGATCTTCTGACAAAGGCAAGTGTCAATTTGAGGAG 834
DB |||||
QY 421 GGAGAGGAAGCAAGGGCTGAATGTTACGTTTATTTGAATAGGCAAAACTGGA 480
DB |||||
QY 835 GGAGAGGAAGCAAGGGCTGAATGTTATTTATTTGAATAGGCAAAAGCTGGA 894
DB |||||
QY 481 TACAATGTAAGCAGCTCTTTTCGAGCTGACAGCAGCTTTTCGCGGGAATCGGAAGCACA 540
DB |||||
QY 895 TACAATGTAAGCAGCTCTTTTCGAGCTGACAGCAGCTTTTCGCGGGAATCGGAAGCACA 954
DB |||||
QY 541 CAGACGGAAGCAGAGAGACATGAGTGACATATAAACTGGAAAGCCTTCAGGACGAACA 600
DB |||||
QY 955 CAGGACGAAGCAGAGAGATATGATGACATATAAACTGGAAAGCCTTCAGGACGAACA 1014
DB |||||
QY 601 GTCAGGAAGGGGTGTTCTCTGCTACTCTCCATGTCATCTCAACCT----- 650
DB |||||
QY 1015 GTCAGGAAGGGGTGTTCTCTGCTACTCTCCATGTCATCTCAACCTCTTCAGAGC 1074
DB |||||
QY 651 --TCTCAGAGAGCCCTTACTCTTTTATTTGACTGCAAGTGTGAATATTTGGCTTGAACCTT 708
DB |||||
QY 1075 TCACTGCTTTGGCCCCCTTACTCTTTTATTTGACTGCAAGTGTGAATATTTGGCTTGAACCTT 1134
DB |||||

QY 709 TTCCCTTCATTAAACGTTTTCGAATTCATCTGCTGCTGTCTCGTGAGA 762
DB |||||
DB 1135 TTCCCTTCAGTAATAACGTTATTCGAATTCATCTGCTGCTGTCTCGTGAGA 1188

RESULT 10

AAZ93836

ID AAZ93836 standard; DNA; 739 BP.

XX AAZ93836;

XX AAZ93836;

DT 29-AUG-2000 (first entry)

XX RAB6C coding sequence.

DE

XX

XX WH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;

KW antibody; immunogen; mutation; detection; therapy; human; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 70..696

FT /*tag= a

FT /label= RAB6C polypeptide

XX W0200029625-A1.

PN

XX 25-MAY-2000.

PD

XX 18-NOV-1999; 99WO-US027630.

PF

XX 18-NOV-1998; 98US-0108994P.

PR

XX

PA (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.

XX

XX Duffy HX, Shan J, Yuan L, Budman D, Calabro A;

XX

XX WPI; 2000-387828/33.

DR P-PSDB; AAY83400.

DR

XX

PS Example 1; Fig 3a-3b; 82pp; English.

XX

XX Nucleic acids encoding WTH3 and RAB6 were isolated by Methyl- (or mutant)
CC Differential Display (MDD) methods and can be used in methods for
CC detecting methylation patterns in multiple drug resistance loci. Genes
CC are frequently not methylated in cells where they are expressed but are
CC methylated in cell types where they are not expressed. Tumour cell DNA is
CC often methylated to a different extent and in different regions when
CC compared to DNA of normal cells. The methylation pattern in a multiple
CC drug resistance locus can be altered and give rise to altered expression
CC patterns of that multiple drug resistance locus. Nucleic acids
CC corresponding to the identified loci such as WTH3 and RAB6 nucleic acids
CC can be used as probes for detecting mutations and methylation patterns of
CC those loci. The nucleic acids and their homologues are useful for
CC inhibition of multiple drug resistance and for treating tumors exhibiting
CC multiple drug resistance. They are also useful for detecting and
CC measuring the expression of mRNA from identified genes and for
CC determining suitable therapeutic treatment. Antibodies directed against
CC immunogenic fragments of WTH3 and RAB6 are useful for detecting specific
CC proteins and polypeptides in tissues or body fluids of patients

SQ Sequence 739 BP; 219 A; 151 C; 191 G; 178 T; 0 U; 0 Other;

Query Match 79.9%; Score 609.2; DB 3; Length 739;

Best Local Similarity 95.6%; Pred. No. 2.6e-179;

Matches 637; Conservative 0; Mismatches 28; Indels 1; Gaps 1;

QY 1 ATGTCGGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 60

Db 70 ATGTCCACGGCGGAGACTTCGGGAATCGCTGAGGAAATTCAGCTGTGTCTCTGGGG 129
Qy 61 GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGATATGACAGTTTTCACAAAC 120
Db 130 GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGATATGACAGTTTTCACAAAC 189
Qy 121 ACTTATCAGGCAATTAATGGCATTGACTTTTATCAAAATCTATGTACTTGGAGGATGGA 180
Db 190 ACTTATCAGGCAATTAATGGCATTGACTTTTATCAAAATCTATGTACTTGGAGGATGGA 249
Qy 181 ACAATCGGGCTTCGGCTGTGGATACGGCGGTGAGGAACGCTTCGTAGCCCTCATTTCCC 240
Db 250 ACAATCAGGCTTCAGCTGTGGGATACGGCGGTGAGGAACGTTTCGTAGCCCTCATTTCCC 309
Qy 241 AGGTACATCCGATTCCTGCTCAGCTGTAGTTAGTTACGATATCACAAATGTTAACTCA 300
Db 310 AGTTACATCCGATTCCTGCTCAGCTGTAGTTAGTTACGATATCACAAATGTTAACTCA 369
Qy 301 TTCCAGCAAACTACAAAGTGGATTGATGTGACAGAAAGAGGAAAGTGTATGTTATC 360
Db 370 TTCCAGCAAACTACAAAGTGGATTGATGTGACAGAAAGAGGAAAGTGTATGTTATC 429
Qy 361 ATCAGCTAGTAGGAAATAGAACAGATCTTGTGACAGAGCAAGTGTCAATTGAGGAG 420
Db 430 ATCAGCTAGTAGGAAATAGAACAGATCTTGTGACAGAGCAAGTGTCAATTGAGGAG 489
Qy 421 GGAGAGAGAAAGCCAAAGGCTGAATGTAGTTTATTTGAACTAGGCAAAACTGGA 480
Db 490 GGAGAGAGAAAGCCAAAGGCTGAATGTATTTTATTTGAACTAGTCAAAAGCTGGA 549
Qy 481 TACAATGTAAGCAGCTCTTCGAGTGTAGCAGCAGCTTTCCCGGAATGGAAGCACA 540
Db 550 TACAATGTAAGCAGCTCTTCGAGTGTAGCAGCAGCTTTCCCGGAATGGAAGCACA 609
Qy 541 CAGGACGGAGCAGAGAACATGATGATGACATATAAACTGGAAGGCTCAGGAGCAAAACA 600
Db 610 CAGGACGAGAGCAGAGAACATGATGATGACATATAAACTGGAAGGCTCAGGAGCAACCA 669
Qy 601 GTCAGGAAAGGGGTGTTCCCTGCTACTCTCCCATGTATCTCAACCCCTTCTCAGAAG 660
Db 670 GTCAGTGAAGAGGCTGTTCCCTGCTAAATCTCCCATGTATCTTCAA-CCTTCTTCAAG 728
Qy 661 CCCCT 666
Db 729 CTCACT 734

RESULT 11
AAZ93841
ID AAZ93841 standard; DNA; 624 BP.
XX
AC AAZ93841;
XX
DT 29-AUG-2000 (first entry)
XX
DE RAB6C coding sequence.
XX
KW WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;
KW antibody; immunogen; mutation; detection; therapy; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..624
FT /*tag= a
FT /product= "RAB6C"
XX
PN WO200029625-A1.
XX
PD 25-MAY-2000.
XX
PF 18-NOV-1999; 99WO-US027630.

XX 18-NOV-1998; 98US-0108994P.
XX (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.
XX Duffy HX, Shan J, Yuan L, Budman D, Calabro A;
XX WPI; 2000-387828/33.
XX P-PSDB; AAY83403.
XX Novel nucleic acids encoding wth3 protein useful for inhibiting multiple
XX drug resistance in tumor cells for identifying the modulators of drug
XX resistance.
XX Claim 6; Page 71-72; 82pp; English.
XX Nucleic acids encoding WTH3 and RAB6 were isolated by Methyl- (or mutant)
XX Differential Display (MDD) methods and can be used in methods for
XX detecting methylation patterns in multiple drug resistance loci. Genes
XX are frequently not methylated in cells where they are expressed but are
XX methylated in cell types where they are not expressed. Tumour cell DNA is
XX often methylated to a different extent and in different regions when
XX compared to DNA of normal cells. The methylation pattern in a multiple
XX drug resistance locus can be altered and give rise to altered expression
XX patterns of that multiple drug resistance locus. Nucleic acids
XX corresponding to the identified loci such as WTH3 and RAB6 nucleic acids
XX can be used as probes for detecting mutations and methylation patterns of
XX those loci. The nucleic acids and their homologues are useful for
XX inhibition of multiple drug resistance and for treating tumors exhibiting
XX multiple drug resistance. They are also useful for detecting and
XX measuring the expression of mRNA from identified genes and for
XX determining suitable therapeutic treatment. Antibodies directed against
XX immunogenic fragments of WTH3 and RAB6 are useful for detecting specific
XX proteins and polypeptides in tissues or body fluids of patients
XX
SQ Sequence 624 BP; 199 A; 115 C; 162 G; 148 T; 0 U; 0 Other;

Query Match 76.9%; Score 585.6; DB 3; Length 624;
Best Local Similarity 96.2%; Pred. No. 5.6e-172;
Matches 600; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGTCGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGTGTCTCTGGGG 60
Db 1 ATGTCGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGTGTCTCTGGGG 60
Qy 61 GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGTATGACAGTTTTCACAAAC 120
Db 61 GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGTATGACAGTTTTCACAAAC 120
Qy 121 ACCTATCAGGCAATTAATGGCATTGACTTTTATCAAAATCTATGTACTTGGAGGATGGA 180
Db 121 ACCTATCAGGCAATTAATGGCATTGACTTTTATCAAAATCTATGTACTTGGAGGATGGA 180
Qy 181 ACAATCGGGCTTCGGCTGTGGATACGGCGGTGAGGAACGCTTCGTAGCCCTCATTTCCC 240
Db 181 ACAATCAGGCTTCAGCTGTGGGATACGGCGGTGAGGAACGTTTCGTAGCCCTCATTTCCC 240
Qy 241 AGGTACATCCGATTCCTGCTCAGCTGTAGTTAGTTACGATATCACAAATGTTAACTCA 300
Db 241 AGTTACATCCGATTCCTGCTCAGCTGTAGTTAGTTACGATATCACAAATGTTAACTCA 300
Qy 301 TTCCAGCAAACTACAAAGTGGATTGATGTGACAGAAAGAGGAAAGTGTATGTTATC 360
Db 301 TTCCAGCAAACTACAAAGTGGATTGATGTGACAGAAAGAGGAAAGTGTATGTTATC 360
Qy 361 ATCAGCTAGTAGGAAATAGAACAGATCTTGTGACAGAGCAAGTGTCAATTGAGGAG 420
Db 361 ATCATGCTAGTAGGAAATAGAACAGATCTTGTGACAGAGCAAGTGTCAATTGAGGAG 420
Qy 421 GGAGAGAGAAAGCCAAAGGCTGAATGTATGTTTATTTGAACTAGGCAAAACTGGA 480
Db 421 GGAGAGAGAAAGCCAAAGGCTGAATGTATGTTTATTTGAACTAGTCAAAAGCTGGA 480

CC patterns of that multiple drug resistance locus. Nucleic acids
CC corresponding to the identified loci such as WH3 and RAB6 nucleic acids
CC can be used as probes for detecting mutations and methylation patterns of
CC those loci. The nucleic acids and their homologues are useful for
CC inhibition of multiple drug resistance and for treating tumors exhibiting
CC multiple drug resistance. They are also useful for detecting and
CC measuring the expression of mRNA from identified genes and for
CC determining suitable therapeutic treatment. Antibodies directed against
CC immunogenic fragments of WH3 and RAB6 are useful for detecting specific
CC proteins and polypeptides in tissues or body fluids of patients
XX
SQ Sequence 573 BP; 162 A; 115 C; 159 G; 137 T; 0 U; 0 Other;

Query Match 69.0%; Score 526; DB 3; Length 573;

Best Local Similarity 100.0%; Pred. No. 2.2e-133;

Matches 526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCGGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 60
DB |||||||
QY 48 ATGTCGGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 107
DB |||||||
QY 61 GAGCAAGCGTTGCCAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGGACAAC 120
DB |||||||
QY 108 GAGCAAGCGTTGCCAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGGACAAC 167
DB |||||||
QY 121 ACCTATCAGGCAATTAATTGGCATTGACCTTTTATCAAAAACATGATCTCGAGGATGGA 180
DB |||||||
QY 168 ACCTATCAGGCAATTAATTGGCATTGACCTTTTATCAAAAACATGATCTCGAGGATGGA 227
DB |||||||
QY 181 ACAATCGGGCTTCGGCTGTGGGATACGGCGGTTCAGGAACGCTCTCCGTAGCTCATTTCCC 240
DB |||||||
QY 228 ACAATCGGGCTTCGGCTGTGGGATACGGCGGTTCAGGAACGCTCTCCGTAGCTCATTTCCC 287
DB |||||||
QY 241 AGGTACATCCGTGATCTCTGTCAGCTGTAGTTTACGATATCACAATAATGTTAACTCA 300
DB |||||||
QY 288 AGGTACATCCGTGATCTCTGTCAGCTGTAGTTTACGATATCACAATAATGTTAACTCA 347
DB |||||||
QY 301 TTCCAGCAAACTTACAAAGTGGATGATGTCAGAACAGAAAGGAGGATGTTATC 360
DB |||||||
QY 348 TTCCAGCAAACTTACAAAGTGGATGATGTCAGAACAGAAAGGAGGATGTTATC 407
DB |||||||
QY 361 ATCAGCTAGTAGGAAATAGAACAGATCTTGTCTGCAAGAGGCAAGTGTCTAGTTGAGGAG 420
DB |||||||
QY 408 ATCAGCTAGTAGGAAATAGAACAGATCTTGTCTGCAAGAGGCAAGTGTCTAGTTGAGGAG 467
DB |||||||
QY 421 GGAGAGGAAAGCCAAAGGGCTGAATGTTACGTTTATGAAACTTAGGGCAAAACTGGA 480
DB |||||||
QY 468 GGAGAGGAAAGCCAAAGGGCTGAATGTTACGTTTATGAAACTTAGGGCAAAACTGGA 527
DB |||||||
QY 481 TACATGTAAGGACGCTCTTTTCGAGCTGACGAGCAGCTTTGCGCG 526
DB |||||||
QY 528 TACATGTAAGGACGCTCTTTTCGAGCTGACGAGCAGCTTTGCGCG 573
DB |||||||

RESULT 15

AAS87692

ID AAS87692 standard; cDNA; 3195 BP.

XX AC AAS87692;

XX AC AAS87692;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #23496.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG23505.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 23496; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 3195 BP; 878 A; 698 C; 699 G; 920 T; 0 U; 0 Other;

Query Match 64.3%; Score 489.8; DB 5; Length 3195;

Best Local Similarity 88.1%; Pred. No. 1.1e-141;

Matches 700; Conservative 0; Mismatches 62; Indels 33; Gaps 14;

QY 1 ATGTCCCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 60
DB |||||||
QY 443 ATGTCCACGGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 502
DB |||||||
QY 61 GAGCAAGCGTTG-CAAAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGGACAA 119
DB |||||||
QY 503 GAGCAAGCGTTG-TGTGAAGACATCTTTGATCACCAGATTGATGATGACAGTTTGGACAA 562
DB |||||||
QY 120 CAGCTATCAGGCAATTAATTCGCAATTTGACTTTTATCAAAACTATGTTACCTGGAGGATGG 179
DB |||||||
QY 563 CAGCTATCAGGCAATTAATTCGCAATTTGACTTTTATCAAAACTATGTTACCTGGAGGATGG 622
DB |||||||
QY 180 AACAAATCGGGCTTCGGCTGTGGGATACGGCGGTTCAGGAACGCTCTCCGTAGCTCATTTCC 239
DB |||||||
QY 623 AACAAATCGGGCTTCGGCTGTGGGATACGGCGGTTCAGGAACGCTCTCCGTAGCTCATTTCC 682
DB |||||||
QY 240 CAGTATCAT-CCGTGATCTCTGTCAGCTGTAGTTTACGATATCACAATAATGTTAACT 298
DB |||||||
QY 683 CAGTTATATCCCGTGATTTCTGTCAGCTGTAGTTTACGATATCACAATAATGTTAACT 742
DB |||||||
QY 299 CATTTCCAGCAAA---CTACAAAGTGGATTCATGATGTCAGAACAGAAAGA-GGAAGTGAT 354
DB |||||||
QY 743 CATTTCCAGGCAAACTACAAAGTGGGATTTGATGTCAGAACAGAAAGGAGGATGAT 802
DB |||||||
QY 355 GTTA-TCATCAGCTAGTAGGAAATAGAACAGATCTTTGCTGACAAAGGCAAGTGTCAGT 413
DB |||||||

Db 803 GTTATTTCATCATGTCTAGTAGGAAATAAAAACAGATCTTGCTGACAAGAGGCAAGTGTCAAT 862
Qy 414 TCAGGAGGGAGAGAGGAAAGCC-AAAGGCTGAATGTT-ACGTTTATTGAAACTAGGCA 471
Db 863 TCAGGAGGGAGAGAGGAAAGCCAAAGAGCTGAATGTTAATTAATGGAACCTAGTGCA 922
Qy 472 AAAACTGG-ATACAATGTAAAGCAGC-TCTTTGACGTGTA--GCAGCAGCTTTGCCGGG 527
Db 923 AAAGCTGGAATACAATGTAAAGCAGCTTCTTCGACGTGTAAGCAGCCAGCTTTGCCGGG 982
Qy 528 AATGAAAGCACACAGGACGGAAGCAGAGAGACATGATGACATAAACTGGAAGGCC 587
Db 983 AATGAAAGCACACAGGACGGAAGCAGAGAGATATGATTGACATAAACTGGAAGGCC 1042
Qy 588 TCAGGAG-CAAAACAGTCAGCGAAGGGG--TTGTTCTGCTACTCTCCCATGTCATCTTC 644
Db 1043 TCAGGAGCCNACCAGTCAGTGAGGAGGCGCTGTTCTGACTAATCTCCCATGTCATCTTC 1102
Qy 645 AACCTTCTCTCAGAA-----GCCCCCTTACTCTTTTCATTGACTGCAGTG 688
Db 1103 AACCTTCTCTTCAGAAAGCTCACTGGCTTTTGGCCCCCTTACTCTTTTCATTGACTGCAGTG 1162
Qy 689 TGAATATTGGCTTGA-ACCTTTTCCCTTCATTAAACGTTTGGCAATTCATCATCTGCTG 747
Db 1163 TGAATATTGGCTTGAATGCGCTTTTCCCTTCAGTAATAACGTAATTGGGATTCATCATCTGCTG 1222
Qy 748 CCTGTCTCGTGGAGA 762
Db 1223 CCTGTCTCGTGGGA 1237

Search completed: April 25, 2005, 09:44:22
Job time : 483 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 07:43:50 ; Search time 177 seconds
(without alignments)
7044.310 Million cell updates/sec

Title: US-09-441-857-11
Perfect score: 762
Sequence: 1 atgtccgcggcgagactt.....tgcgtcgtctcgtggaga 762

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution..

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	639.2	83.9	3745	US-09-976-594-387	Sequence 387, App
2	564.4	74.1	739	US-09-949-016-2946	Sequence 2946, App
3	204	26.8	329	US-09-513-999C-2679	Sequence 2679, App
4	152.4	20.0	86857	US-09-949-016-14688	Sequence 14688, A
5	135.6	17.8	414	US-09-248-796A-6213	Sequence 6213, App
6	117.8	15.5	1631	US-09-620-312D-587	Sequence 587, App
7	116.2	15.2	719	US-09-949-016-3023	Sequence 3023, App
8	116.2	15.2	1984	US-09-023-655-7	Sequence 7, Appli
9	111.8	14.7	921	US-09-016-434-1124	Sequence 1124, App
10	111	14.6	1546	US-09-949-016-5540	Sequence 5540, App
11	105.6	13.9	24257	US-09-949-016-13902	Sequence 13902, A
12	102.2	13.4	848	US-08-741-411-2	Sequence 2, Appli
13	101.4	13.3	1255	US-09-949-016-1772	Sequence 1772, App
14	99.8	13.1	8137	US-09-566-921-7	Sequence 7, Appli
15	99.8	13.1	723	US-09-016-434-1422	Sequence 1422, App
16	99.8	13.1	920	US-09-949-016-4287	Sequence 4287, App
17	99.4	13.0	833	US-09-620-312D-426	Sequence 426, App
18	98.8	13.0	639	US-09-399-913-66	Sequence 66, Appli
19	98.8	13.0	639	US-09-350-614-66	Sequence 66, Appli
20	98.8	13.0	1148	US-09-949-016-4879	Sequence 4879, App
21	94	12.3	1775	US-09-949-016-4926	Sequence 4926, App
22	92.8	12.2	1069	US-09-620-312D-646	Sequence 646, App
23	92	12.1	642	US-09-248-796A-6190	Sequence 6190, App
24	91.8	12.0	1308	US-09-270-767-12890	Sequence 12890, A
25	90.8	11.9	1090	US-09-799-451-914	Sequence 914, App
26	90.6	11.9	601	US-09-949-016-73802	Sequence 73802, A
27	90.6	11.9	1630	US-09-949-016-409	Sequence 409, App

ALIGNMENTS

RESULT 1

US-09-976-594-387
; Sequence 387, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 387
; LENGTH: 3745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 107569.15
US-09-976-594-387

Query Match	83.9%	Score 639.2;	DB 4;	Length 3745;
Best Local Similarity	90.3%	Pred. No. 7.2e-202;		
Matches	699;	Conservative	0;	Mismatches 63;
Indels	12;	Gaps	1;	
Qy	1	ATGTCGGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG	60	
Db	494	ATGTCACGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG	553	
Qy	61	GAGCAACCGTTCGAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTTCACAAAC	120	
Db	554	GAGCAACCGTTCGAAGACATCTTGTATCACCAGATTCAGGTATGACAGTTTTCACAAAC	613	
Qy	121	ACCTATCAGGCAATTAATGGCAATTCACATTTTATCAAAACTATGTACTTGGAGGATGGA	180	
Db	614	ACCTATCAGGCAACAAATTTGGCAATTCACATTTTATCAAAACTATGTACTTGGAGGATGGA	673	
Qy	181	ACAATCGGGCTTCGGCTGTGGGATACGGCGGTCAGGACGTCCTCCGTAGCTCATTTCCC	240	
Db	674	ACAGTACGATTCGAATTTATGGGACACACAGCGTTCAGGAGCTTGTGATTCCT	733	
Qy	241	AGGTACATCCGCTGATTCGTCGAGCTGTAGTATTTACGATATACACAAATGTTAACTCA	300	
Db	734	AGGTACATTCGCTGATTCGTCGAGCTGTAGTATTTACGATATACACAAATGTTAACTCA	793	
Qy	301	TTCCAGCAAACTACAAAGTGGATTCATGATCTCAGAACAGAAAGGAGGTGATGTTATC	360	


```
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 587
; LENGTH: 1631
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (66)..(650)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1631)
; OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-587
```

```
Query Match 15.5%; Score 117.8; DB 4; Length 1631;
Best Local Similarity 53.1%; Pred. No. 2.6e-28;
Matches 251; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

Qy 29 CGCTGAGGAATTCAAGCTGGTTCCTGGGGGAGCAAGCGTTGCCAAGACATCTTTGA 88
Db 70 CGCTGAGGAGCTCAAGATGTCTCTGCTGGGGATACAGGTGTAGTAAATCGAGTATTG 129

Qy 89 TCACAGAGATTGAGTATGACAGTTGTTTGCAGCAATATTCAGGCAATATTCGCAATGACT 148
Db 130 TGTGCGGTTTGTGGAAGACAGTTTGTATGATCAAAACATCAACCAACAATAGGGGATCTT 189

Qy 149 TTTTATCAAAACTATGTACTTGGAGGATGGAACAATCGGGTTCGGCTGTTGGGATACGG 208
Db 190 TTATGACCAAGACTGCTCCAGTACCAAAATGAGCTACATAAAATTCCTAATCTGGGATACAG 249

Qy 209 CGGGTCAGGAAGCTCTCGTAGCTCATTCCAGGTACATCGTGATTCTGCTGCAGCTG 268
Db 250 CTGGAACAAGACGATTTCTGCTTAGCACCAGTATCTATCGAGGGTGGCTGCAGCTA 309

Qy 269 TAGTAGTTTACGATATCAAAATGTTAACTCATTCAGCAAACTACAAAGTGGATTGATG 328
Db 310 TAATCGTTTATGATATCAAAAGAAGAGACATTTTCAACATTAAGAATTTGGGTGAAAG 369

Qy 329 ATGTCAGAACAGAAAGAGAGTGTATCATCACCTAGTACCTAGTAAAGTAAAGACATC 388
Db 370 AGCTTCGACAGCATGGCCCACTTAATATTGTAGTTGCCATTGCGAGGAAATAAATGTGATC 429

Qy 389 TTGCTGCAAGAGCAAGTGTGAGTTGAGGGAGGAGAGGAAAGCCAAAGGGCTGAATG 448
Db 430 TTATCGATGTAGAGAGTGTGAGAGAGATGCAAGGACTACGCCGACTCTATTTCATG 489

Qy 449 TTACGTTTATTGAACTAGGGCAAAACTGGATACAATGTAAAGCAGCTCTTT 501
Db 490 CAATTTTGTAGAGACCGCGCAAAACCGGATAAATCAATGAACTCTTT 542
```

```
RESULT 7
US-09-949-016-3023
; Sequence 3023, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3023
; LENGTH: 719
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3023
```

```
Query Match 15.2%; Score 116.2; DB 4; Length 719;
Best Local Similarity 51.7%; Pred. No. 5.2e-28;
Matches 265; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

Qy 22 GGGAAATCGGTGAGGAATTCAGCTGGTTCCTGGGGGAGCAAGCGTTGCCAAGACA 81
Db 99 GGAATAATAATATGCCAGTTCAAACTAGTACTTCTGGGAGAGTCCGCTGTTGGCAATCA 158

Qy 82 TCTTTGATCACCAAGATTCAGGTATGACAGTTTTCACAAACACCTATCAGGCAATATTTGGC 141
Db 159 AGCCTAGTGTCTGTTTGTGAAAGGCCAATTCATGAATTTCAAGAGAGTACCATTTGGG 218

Qy 142 ATTGACTTTTTATCAAAACTATGTACTTGGAGGATGGAACAATCGGGCTTCGGCTGTGG 201
Db 219 GCTGCTTTTCTAACCCCAACTGTATGTCTTGATGACACTACAGTAAAGTTTGAATAATGG 278

Qy 202 GATACGGGGGTGAGGAACGCTCCGTAGCCTCATTCAGGCTACATCCGCTGATCTTGCT 261
Db 279 GATACAGCTGGTCAAGAACGATACCATAGCCTAGCACCATGTACTACAGAGGAGCAAA 338

Qy 262 GCAGCTGTAGTGTATGATATCACAAAATGTTAACTCATTCAGCAAACTACAAAGTGG 321
Db 339 GCAGCCATGTGTATGATATCACAAAATGAGAGTCTCTTGGCAAGGCAAAATTTGG 398

Qy 322 ATTGATGATGTGACAAAGAGAGAGTGTATCATCACGCTAGTAGGAAATAGA 381
Db 399 GTTAAAGAACTTCAGAGGCAAGCAAGTCTTAACATTGTAATAGCTTTTATCGGAAACAAG 458

Qy 382 ACAGATCTTCTGACAGAGGCAAGTGTCTAGTTGAGGGAGGAGAGAGAAAGCCAAAGGG 441
Db 459 GCCAGCTAGCAATAAAGAGCAGTAGATTCCAGGAAGCAGAGTCTCTATGCAAGATGAC 518

Qy 442 CTGAATCTTACGTTTATTGTTGAACTTAGGGCAAAACTGGATACAAATGTAAAGCAGCTCTTT 501
Db 519 AATAGTTTATTATTCATGAGAGACATCCGCTTAAACATCAATGAATGTAAATGAATATTC 578

Qy 502 CGAGCTGTAGCAGCAGCTTTGCGGGGAATGGAA 534
Db 579 ATGGCAATAGCTAAAAAATTTGCCAAAGAAATGAA 611
```

```
RESULT 8
US-09-023-655-7
; Sequence 7, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1984 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: U937NOT01
CLONE: 000513
US-09-023-655-7

Query Match 15.2%; Score 116.2; DB 4; Length 1984;
Best Local Similarity 51.7%; Pred. No. 1e-27;
Matches 265; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 22 GGAATCCGCTGAGGAATTCAGCTGGTTCCTGGGGGAGCAAGCGTTGCAAGACA 81
DB 11 GGAATAAAATATGCCAGTTCAAAGTAGTACTCTGGGAGAGTCCGCTGTGGCAATCA 70
QY 82 TCTTTGATCACCAGATTCAGGTATGACGTTTGTGACACACCTATCAGGCATATTCG 141
DB 71 AGCCTAGTCTCGTTTGTGAAAGGCCAATTCATGAATTTCAAGAGAGTACCAATGGG 130
QY 142 ATTGACTTTTATCAAAACTATGTACTTGGAGGATGGAACAATCGGGCTTCGGCTGTGG 201
DB 131 GCTGCTTTTCTAACCCAACTGTATGTCTGTATGACACTACAGTAAAGTTTGAATATGG 190
QY 202 GATACGGCGGTGAGGAAGCTCTCGTAGCCTCATTTCCAGGTATACATCCGTTGATTCGT 261
DB 191 GATACAGCTGGTCAAGAACGATACCATAGCTAGCACCACCAATGTACTACAGAGGAGCACA 250
QY 262 GCAGCTGTAGTAGTTACATATACAAATGTTAACTCATTCAGCAAACTACAAAGTGG 321
DB 251 GCAGCCATAGTTGTATATGATATACAAATGAGGAGTCTCTTTGCAAGAGCAAAAATGG 310
QY 322 ATTGATGATGTACAGACAGAAAGAGGAGTGTATGTTATCATCAGCTAGTAGGAATAGA 381
DB 311 GTTAAAGAACTTCAGAGGCAAGCAAGTCTTAACATTTGTAATAGCTTTATCGGGAACAG 370
QY 382 ACAGATCTTCTGACAAAGGCAAGTGTGAGTGGAGGGAGAGAGGAAAGCCAAAGG 441
DB 371 GCGACCTAGCAAAATAAAGAGCAGTAGATTTCCAGGAAGCAGAGTCTCTATGCAATGAC 430
QY 442 CTGAATGTTACGTTTATGAACTAGGCAAAAACCTGGATACATGTAAGCAGCTCTTT 501
DB 431 AATAGTGTATTTATTCATGAGAGCATCCGCTAAACATCAATGAATGAAATGAAATATTC 490
QY 502 CGACGTTGAGCAGCAGCTTTTCCGGGAATGGAA 534
DB 491 ATGGCAATAGTAAAAAATTTGCCAAAGAAATGAA 523

RESULT 9

US-09-016-434-1124
Sequence 1124, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1124:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1388194
US-09-016-434-1124

Query Match 14.7%; Score 111.8; DB 4; Length 921;
Best Local Similarity 51.5%; Pred. No. 1.8e-26;
Matches 257; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

QY 8 CGGGGGGAGACTTCGGGAATCCGCTGAGGAATTCAGCTGGTGTCTCTGGGGGAGCAA 67
DB 44 CTGGCTGCGAGCACATGATGCGGATACCGGAGAGCTCAAAGTGTGCTTCTCGGGGACATG 103
QY 68 GGGTTGCAAGACATCTTTTGATCACCAGATTCAGGTATGACAGTTCGATTTTGACACACCTATC 127
DB 104 GGGTTGGAATCAAGCATCGTGTGCTGATTTGTCAGGATCATTTCACCAACATCA 163
QY 128 AGGCAATAATTGGCATTCGACTTTTATCAAAAACCTATGTACTTTGAGGATGGAACAATCG 187
DB 164 GCGCTACTATTGGGGCATCTTTTATGACCAAACTGTGCTTGTGGAATGAATTCACA 223
QY 188 GCGTTGCGCTGTGGGATACGGGGCTCAGGAACGCTCTCGTAGCCTCATTTCCAGGTACA 247
DB 224 AGTTCTCATCTGGGACACTGCTGCTCAGGAACGGTTTCATTTCATTGGCTCCCATGTACT 283
QY 248 TCCGTGATTTGCTGCAGCTGTAGTGTATGATATCAGATATCAGAAATGTTAACTCATTCACG 307
DB 284 ATCGAGGCTCAGCTCAGCTGTTATCGTGTATGATATTTACCAGCAGGATTCATTTTATA 343
QY 308 AAACTACAAAGTGGATTGATGATGTGACAGAAAGAGGAGTGTATGTTATCATCACGC 367
DB 344 CCTTGAAGAAATGGCTCAAGGAGCTGAAAGAACATGGTCCAGAAACATTTGTAATGGCCA 403
QY 368 TAGTAGGAAATAGAACAGATCTTGTGTGACAGAGGCAAGTGTGAGTTGAGGAGGAGAGA 427
DB 404 TCGCTGGAACAAGTGCAGCTCTCAGATATTTAGGAGGTTCCCTCTGAAGGATGCTTAAGG 463

```
Qy 428 GGAAGCCAAAGGGCTGAATGTTACGTTTATTGAACCTAGGCAAAAACCTGGATACAATG 487
Db 464 BATAAGCTGAATCCATAGTGCCTCGTGGTTGACAAAGTGCAAAAATGCTATTATA 523

Qy 488 TAAAGCAGCTCTTTCGACG 506
Db 524 TCGAAGAGCTCTTTCAAGG 542

RESULT 10
US-09-949-016-5540
; Sequence 5540, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5540
; LENGTH: 1546
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5540

Query Match 14.6%; Score 111; DB 4; Length 1546;
Best Local Similarity 52.0%; Pred. No. 4.6e-26;
Matches 249; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy 22 GGAATCCGCTGAGGAATTCAGCTGTGTTCTGGGGGAGCAAGGTTGCAAGACA 81
Db 134 GGAACAAGATCTGTCAATTTAAAGCTGTTCTGCTGGGGGAGTCTGCGGTAGGCAATCC 193

Qy 82 TCTTTGATCACCAGATTCAGGTATGACAGTTTGTGACAACTATCAGGCAATATTTGGC 141
Db 194 AGCCTGCTCTCCGCTTTGTCAAGGACAGTTTACAGAGTACCAGGAGACAAATGGA 253

Qy 142 ATTGACTTTTATCAAAAACATATGTAATTGAGGATGGAACAATCGGGCTTCGGCTGTGG 201
Db 254 GCGGCTTTCTCCACACAGACTGTCTGCTGGATGACACAACAGTCAAGTTTGAGATCTGG 313

Qy 202 GATACGGGGGTGAGGAAGCTCTCGGTAGCCTCATTCACAGGTACATCCGTGATTCGTCT 261
Db 314 GACACAGCTGACAGGAGCGGTATCACAGCTTGGCCCCCATGTATTCGGGGGGGCCGAG 373

Qy 262 GCAGCTGTAGTAGTTTACGATATCAAAATGTTAACTCATTCAGCAAACTACAAAGTGG 321
Db 374 GCTGCCATCGTGCTATGATCATCCAAACACAGTACATTTGACGGGGCAAGACTGG 433

Qy 322 ATTGATGATGTCAGAACAGAAGAGGAAGTATGTTATCATCAGCTAGTAGGAAATAGA 381
Db 434 GTGAAGGAGCTACAGAGGACGCCAGCCCCAACATCGTCAATTCGCACTCGCGGGTAAACAAG 493

Qy 382 ACAGATCTTGTGACAAAGGCAAGTGTCAAGTTGAGGAGGAGAGAGAAAGCCAAAGGG 441
Db 494 GCAGACCTGGCCAGCAAGAGAGCGGTGGAATTTCCAGGAAGCAGCAAGCCTATGTCAGACGAC 553

Qy 442 CTGAATGTTACGTTTATTGAACCTAGGCAAAAACCTGGATACAATGTTAAAGCAGCTCTT 500
Db 554 AACAGTTTGTCTTCATCGAGACATCAGCAAAAGACTGCAATGAACGTGAACGAATCTT 612

RESULT 11
US-09-949-016-13902/c
```

```
; Sequence 13902, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13902
; LENGTH: 24257
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13902

Query Match 13.9%; Score 105.6; DB 4; Length 24257;
Best Local Similarity 51.7%; Pred. No. 1.7e-23;
Matches 240; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

Qy 40 TTCAGCTGTGTTCCTGGGGGAGCAAGCGTTGCAAAAGACATCTTTGATCACCAGATTC 99
Db 8910 TTCAAGTTGTGCTGTGATCGGGGACTCGGGGGTGGGCAAGACTTGTCTGTGATCATTCGCTTT 8851

Qy 100 AGGTATGACAGATTTTGACAAACCTATCAGGCAATATTCGCAATGACTTTTATCAAAA 159
Db 8950 GCAGAGGACAACTTCAACAAACCTTACATCTCCACCATCGGAATTTTCAAGATCCGC 8791

Qy 160 ACTATGTACTTTGGAGGATGGAACAATCGGGCTTCGGCTGTGGGATACGGGGGTCAAGAA 219
Db 8790 ACTGTGATATAGGGGGAAGAGATCAAACTACAAGTCTGGGACACAGCTGGCCAAGAG 8731

Qy 220 CGTCTCGTAGCTTCATCCAGGTACATCCGTGATTTCTGCTGACGCTGTAGTAGTTTAC 279
Db 8730 CGGTTCAAGACAATACTACTGCTTACTACCGTGGAGCCATGGGCATTTATCTAGCATAC 8671

Qy 280 GATATCAAAATGTTAACTCATTTCCAGCAAACTACAAGTGGATTTGATGATGTCAGAA 339
Db 8670 GACATCCGATGAGAAATCTTTGAGAAATTTTCAGAACTTGGATGAAAGCATCTGAG 8611

Qy 340 GAAAGAGGAAGTGTATTTATCATCAGCTAGTAGGAAATAGAACAGATCTTGTCTGACAAG 399
Db 8610 AATGCCCTCAGCTGGGGTGGAGCACCTCTTCTAGGGAACAAATGTGACATGGAGGCCAAG 8551

Qy 400 AGGCAAGTGTGAGTGGAGGAGGAGAGAAAGCCAAAGGCTGAATGTTACGTTTATT 459
Db 8550 AGGAAGTGTGAGAGGAGGAGCGGATAGTTGGCTCGAGAGCATGGAATCCGATTTTTC 8491

Qy 460 GAAACTAGGCAAAAACCTGGATACAATGTAAGCAGCTCTTTTCG 503
Db 8490 GAAACTAGTCTAAATCCAGTATGAATGTGATGAGGCTTTTAC 8447

RESULT 12
US-08-741-411-2
; Sequence 2, Application US/08741411
; Patent No. 6124116
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: NOVEL RAB PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
```



```
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/741,411
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0139 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 848 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-08-741-411-2

Query Match 13.4%; Score 102.2; DB 3; Length 848;
Best Local Similarity 51.7%; Pred. No. 2.7e-23;
Matches 233; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 56 TGGGGGAGCAAGCGTTGCAAGACATCTTTCATCACCAGATTCAGGTATGACAGTTTG 115
DB 211 TGGAGGACACTGGGGTTGGAAATCAAGCATCGTGTGCAATTTGTCCAGGATCACTTTG 270
QY 116 ACAACACCTATCAGGCAATAATGGCAATGATCTTTTATCAAAAATATGTTACTTGGAGG 175
DB 271 ACACACATCAGCCCTACTATTTGGGCATCTTTATGACCAAACTGTGCCTTTGGGA 330
QY 176 ATGGAACTTCAGAACTTCCTCATCTGGGACACTGCTGTCAGGAACGGTTTCATTCTGG 390
DB 331 ATGAACTTCAGAACTTCCTCATCTGGGACACTGCTGTCAGGAACGGTTTCATTCTGG 390
QY 236 TTCCAGGTACATCGTGTATCTGTCGAGCTGTAGTAGTTTACGATATCACAATGTGA 295
DB 391 CTCCTATGATATCGAGGCTCAGCTGCAGCTGTTATCGTGTATGATATTTACCAAGCAGG 450
QY 296 ACTCATTCAGCAAACTACAAAGTGGATTGATGTCAGACAGAAAGAGGAAGTATG 355
DB 451 ATTCATTTTATACCTTGAAGAAATGGTCAAGAGCTGAAAGAACTATGTCAGAAACA 510
QY 356 TTATCATCATCGGTAGTAGGAATAGAACAGATCTTGTCTGCAAGAGGCAAGTGTGAGTTG 415
DB 511 TTGTAATGCCATCGCTGGAACAAGTGGGACCTCTCATATATTAGGAGGTTCCCTGA 570
QY 416 AGGAGGAGAGAGGAAGCAAGGCTGAATGTTAGCTTTTATGAACTAGGCAAAA 475
DB 571 AGGATGCTAAGGAATACGCTGAATCCATAGGTGCCATCGTGTGTGAGCAAGTGCAGAAA 630
QY 476 CTGGATACAAATGTAAGACGAGCTCTTTCGACG 506
DB 631 ATGCTATTAATTCGAGAGAGCTCTTTCAGG 661

RESULT 13

US-09-949-016-1772
Sequence 1772, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1772
LENGTH: 1255
TYPE: DNA
ORGANISM: Human
US-09-949-016-1772

Query Match 13.3%; Score 101.4; DB 4; Length 1255;
Best Local Similarity 50.7%; Pred. No. 6.4e-23;
Matches 243; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

QY 39 ATTCAGCTGTGTTCTGGGGAGCAAGCGTTGCAAGACATCTTTGATCACCAGATT 98
DB 129 ATTCAGTTACTTCTGATTTGGCGACTCAGGGGTTGGAAAGTCTTTCCTTAGGTT 188
QY 99 CAGGTATGACAGATTTTGAACAACCTATCAGGCAATAATTCGCTATGACTTTTATCAAA 158
DB 189 TGCAGATACATATACAGAAAGTACATCAGACATTTGGTGGATTTCAAAATAG 248
QY 159 AACTATGATCTTGGAGGATGAACAATCGGCTTCGGCTGTGGGATACGGGGGTGAGA 218
DB 249 AACTATAGTTAGACGGGAAAAACAATCAAGCTTCAAAATATGGGACACAGCAGGCAGGA 308
QY 219 ACGTCTCGTAGCTCATCTCCAGGTACATCCGTGATTTCTGCTCAGCTGTAGTAGTTTA 278
DB 309 AAGATTTCCGAACAATCACCTCCAGTTATTACAGAGGAGCCCATGGCATCATAGTTGTGA 368
QY 279 CGATATCAAAATGTTAACTCATTTCCAGCAAACTACAAAGTGGATTTGATGTGAGAAC 338
DB 369 TGATGTGACATCAGGAGTCTTCAATAATGTTAAACAGTGGCTGCAGGAAATAGATCG 428
QY 339 AGAAAGAGGAAGTGTATATCATCACGCTAGTAGGAAATAGAACAGATCTTGTGACAA 398
DB 429 TTATGCCAGTGAAATATGTCAACAAATTTGTTGGTAGGAAACAAATGTGATCTGACCACAA 488
QY 399 GAGGCAAGTGTGATTTAGGAGGAGAGAGAAAGCCAAAGGCTGAATGTTACCTTTAT 458
DB 489 GAAAGTAGTAGACTACACAAACAGCGAAGGAATTTGCTGATTTCCCTTGGAAATTCCTTTT 548
QY 459 TGAAGTAGGCAAAACTGGATACAATGTTAAAGCAGCTCTTTCGACGCTAGCAGCAG 517
DB 549 GGAACACAGTGTAGAAATGCAAGATGTAGAACAGTCTTTTCATGACGATGGCAGCTG 607

RESULT 14
US-09-566-921-7
Sequence 7, Application US/09566921
Patent No. 682888
GENERAL INFORMATION:
APPLICANT: Loring, Jeanne F.
APPLICANT: Tingley, Debora W.
APPLICANT: Edwards, Carla M.
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
FILE REFERENCE: PA-0024 US
CURRENT APPLICATION NUMBER: US/09/566,921
CURRENT FILING DATE: 2000-05-05
```

```
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 8137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6682888 411474.10
; NAME/KEY: unsure
; LOCATION: 3488-3788
; OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-7

Query Match      13.3%; Score 101.4; DB 4; Length 8137;
Best Local Similarity 50.7%; Pred. No. 2.1e-22;
Matches 243; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

Qy 39 ATTCAAGCTGGTTCCTCGGGGAGCAAGCGTTGCAAGACATCTTTGATCACCAGATT 98
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 235 ATTCAAGTTACTTCTGATTGGGACTCAGGGGTTGGAAAGTCTTGCCCTTCTTTAGGTT 294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 99 CAGGTATGACAGTCTTGACACACCTATCAGGCAATATTCGCAATGTCATCTTTTATCAA 158
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 295 TGCAGATGATACATATACAGAAAGCTACATCAGCACAAATGGTGGATTTCAAATAAG 354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 159 AACTATGTACTTGGAGGATGGAACAATCGGCTTGGCTGTGGGATACGGCGGTGAGGA 218
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 355 AACTATAGATTAGACGGGAAACAATCAAGCTTCAAAATATGGGACACAGCAGGCCAGGA 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 219 ACGTCTCGTAGCTTATCCAGGTATACCGGTATCTGCTGAGCTGTAGTAGTTTA 278
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 415 AAGATTTGCAACAATCACCTCCAGTTATTACAGAGGAGCCCATGGCATCATAGTTGTGA 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 279 CGATATCAAAATGTTAACTCTATTCAGCAAACTACAAAGTGGATTTGATGATGTCAGAAC 338
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 475 TGATGTGACATCAGGAGTCTTCAATAATGTTAAACAGTGGCTGCAGGAATAGATCG 534
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 339 AGAAAGAGGAAGTGTATCATCACGCTAGTAGGAAATAGAACAGATCTTGCTGACAA 398
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 535 TTATGCCAGTGAAATGTCAACAAATTTGTTGGTAGGGAACAAATGTGATCTGACCACAA 594
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 399 GAGGCAAGTGTAGTTGAGGAGGAGAGAGAAAGCCAAAGGCTGATGTACGTTTAT 458
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 595 GAAAGTAGTAGATACACACAGCGAAGGAATTTGCTGATTCCTTGGAAATTCGGTTTTT 654
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 459 TGAAGTGGGCAAAATCGATACAAATGTAAGCAGCTCTTTCCAGCTGTAGCAGCAG 517
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 655 GGAACCAAGTCTAAGATGCAACGATGTAGAACAGTCTTTCATGACGATGGCAGCTG 713
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 15

```
US-09-016-434-1422
; Sequence 1422, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1422:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9550059
US-09-016-434-1422
```

```
Query Match      13.1%; Score 99.8; DB 4; Length 723;
Best Local Similarity 50.5%; Pred. No. 1.5e-22;
Matches 242; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

Qy 39 ATTCAAGCTGGTTCCTCGGGGAGCAAGCGTTGCAAGACATCTTTGATCACCAGATT 98
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 83 ATTCAAGTTACTTCTGATTGGGACTCAGGGGTTGGAAAGTCTTGCCCTTCTTTAGGTT 142
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 99 CAGGTATGACAGTCTTGACAAACACCTATCAGGCAATATTTGGGATTCACATCTTTTATCAA 158
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 143 TGCAGATGATACATATACAGAAAGCTACATCAGCACAAATTTGGTGTGATTTCAAATAAG 202
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 159 AACTATGTACTTGGAGGATGGAACAATCGGCTTGGGCTGTGGGATACGGCGGTGAGGA 218
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 203 AACTATAGAGTTAGACGGGAAACAATCAAGCTTCAAAATATGGGACACAGCAGGCCAGGA 262
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 219 ACGTCTCGTAGCTCTATTCAGGATACATCCGTTGATTCGCTGACGCTGTAGTAGTTTA 278
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 263 AAGATTCGAACAATCACCTCCAGTTATTACAGAGGAGCCCATGGCATCATAGTTGTGA 322
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 279 CGATATCAAAATGTTAACTCTATTCAGCAAACTACAAAGTGGATTTGATGATGTCAGAAC 338
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 323 TGATGTGACAGATCAGGAGTCTTCAATAATGTTAAACAGTGGCTGCAGGAATAGATCG 382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 339 AGAAAGAGGAAGTGTATCATCAGCTAGTAGGAAATAGAACAGATCTTGCTGACAA 398
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 383 TTATGCCAGTGAAATGTCAACAAATTTGTTGGTAGGGAACAAATGTGATCTGACCACAA 442
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 399 GAGGCAAGTGTAGTTGAGGAGGAGAGAAAGCCAAAGGCTGATCTTTAGGTTTAT 458
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 443 GAAAGTAGTAGACTACACAAACAGCGAAGGAATTTGCTGATTCCTTGGAAATTCGGTTTT 502
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 459 TGAAGTGGGCAAAATCGATACAAATGTAAGAGCAGCTCTTTCAGCTGTAGCAGCAG 517
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 503 GGAACCAAGTCTAAGATGCAACGACGATAGAACAGTCTTTCATGACGATGGCAGCTG 561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: April 25, 2005, 11:43:33
Job time : 179 secs


```
; NAME/KEY: CDS
; LOCATION: (27078)..(27094)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (141192)..(141769)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159571)..(159606)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (174525)..(174575)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (234891)..(235013)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (235514)..(235560)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (279677)..(279729)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (408660)..(409123)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409204)..(409669)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (428381)..(428396)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472204)..(472330)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (714252)..(714355)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (714447)..(714529)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (739794)..(739891)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (744484)..(744602)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (51812)..(51911)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (57122)..(57221)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (79368)..(79467)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (293951)..(294050)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (310089)..(310188)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (332935)..(332935)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (332992)..(332992)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
```

```
; NAME/KEY: modified_base
; LOCATION: (362002)..(362101)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (639781)..(639880)
; OTHER INFORMATION: a, t, c, g, unknown or other
US-10-292-798-1369

Query Match          99.7%; Score 759.4; DB 17; Length 744802;
Best Local Similarity 99.9%; Pred. No. 6.2e-227;
Matches 760; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTCGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTTCCTCGGG 60
Db 345708 ATGTCGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTTCCTCGGG 345649
Qy 61 GAGCAAAAGCGTTGCAAAAGACATCTTTGATCACCAGATTTCAGGTATGACAGTTTTCACAAC 120
Db 345648 GAGCAAAAGCGTTGCAAAAGACATCTTTGATCACCAGATTTCAGGTATGACAGTTTTCACAAC 345589
Qy 121 ACCTATCAGGCAATAATTGGCATTGACTTTTATCAAAAATACTATGTACTTTGGAGATGGA 180
Db 345588 ACCTATCAGGCAATAATTGGCATTGACTTTTATCAAAAATACTATGTACTTTGGAGATGGA 345529
Qy 181 ACAATCGGGCTTCGGCTGTGGGATACGGCGGGTCAGGAACGCTCCGTAGGCTCATTTCCC 240
Db 345528 ACAATCGGGCTTCGGCTGTGGGATACGGCGGGTCAGGAACGCTCCGTAGGCTCATTTCCC 345469
Qy 241 AGGTACATCCGTGATTCGTCTGAGCTGTAGTATTACGATATACAAATGTTAACTCA 300
Db 345468 AGGTACATCCGTGATTCGTCTGAGCTGTAGTATTACGATATACAAATGTTAACTCA 345409
Qy 301 TTCAGCAAACTACAAAGTGGATTGATGTTCAGAACAGAAAGAGAGATGATGTTATC 360
Db 345408 TTCAGCAAACTACAAAGTGGATTGATGTTCAGAACAGAAAGAGAGATGATGTTATC 345349
Qy 361 ATCAGCTAGTAGGAATAGAACAGATCTTCTGACAGAGGCAAGTGTCAAGTTGAGGAG 420
Db 345348 ATCAGCTAGTAGGAATAGAACAGATCTTCTGACAGAGGCAAGTGTCAAGTTGAGGAG 345289
Qy 421 GGAGAGAGAAAGCCAAAGGCTGAATGTTACGTTTATTGAAACTAGGCGCAAAACTGGA 480
Db 345288 GGAGAGAGAAAGCCAAAGGCTGAATGTTACGTTTATTGAAACTAGGCGCAAAAGCTGGA 345229
Qy 481 TACAATGTAAGCAGCTCTTTTCGACGCTGTAGCAGCAGCTTTGCGGGAATGGAAGCACA 540
Db 345228 TACAATGTAAGCAGCTCTTTTCGACGCTGTAGCAGCAGCTTTGCGGGAATGGAAGCACA 345169
Qy 541 CAGGACGGAAGCAGAGACATGATGACATAAACTGGAAGGCAAGCTCAGGAGCAAAACA 600
Db 345168 CAGGACGGAAGCAGAGACATGATGACATAAACTGGAAGGCAAGCTCAGGAGCAAAACA 345109
Qy 601 GTCAGCAAGGCGGTGTTCTCTACTCTCCATGTCATCTTCAACCCCTTCCTCAGAAG 660
Db 345108 GTCAGCAAGGCGGTGTTCTCTACTCTCCATGTCATCTTCAACCCCTTCCTCAGAAG 345049
Qy 661 CCCCTTACTCTTTTCATTTGACTGCAGTGTGAATATTTGGCTTGAAACCTTTTCCCTTCATTA 720
Db 345048 CCCCTTACTCTTTTCATTTGACTGCAGTGTGAATATTTGGCTTGAAACCTTTTCCCTTCATTA 344989
Qy 721 ATAACGTTTTGCAATTCATCATTTGCTGCCTGTCTCGTGGAG 761
Db 344988 ATAACGTTTTGCAATTCATCATTTGCTGCCTGTCTCGTGGAG 344948

RESULT 2
US-09-925-300-424
; Sequence 424, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
```

Query Match	73.4%	Score 559.6	DB 17	Length 740
Best Local Similarity	91.0%	Pred. No. 1.4e-165		
Matches 606	Conservative	0	Mismatches 59	Indels 1
Qy	1	ATGTCCGCGGCGGAGACTTTCGGGAATCCGCTCAGGAAATTCGAAGCTGGTGTTCCTCGGG	60	
Db	71	ATGTCCACGGCGGAGACTTCGGGAATCCGCTCAGGAAATTCGAAGCTGGTGTTCCTCGGG	130	
Qy	61	GAGCAAGCGTTGCAAGAGACATCTTTGATCACAGATTTCAGGTATGACAGTTTGTGACAAC	120	
Db	131	GAGCAAGCGTTGCAAGAGACATCTTTGATCACAGATTTCATGATGATGACAGTTTGTGACAAC	190	
Qy	121	ACCTATCAGGCAATTAATTTGGCAATTCCTTTTATCAAAAACTATGTACTTTGGAGGATGGA	180	
Db	191	ACCTATCAGGCAACNAATGSCATTGACTTTTATCAAAAATCTATGTACTTTGGAGGATCGA	250	
Qy	181	ACAATCGGGCTTCGGCTGTGGGATACGGCGGTCAGGAACGTCTCCGTAGCCTCATTTCCC	240	
Db	251	ACAGTACGATTGCAATTTATGGGACACAGCAGGTCAAGAGCGGTTTCAGAGAGCTTGATTTCT	310	
Qy	241	AGGTACATCCGTGATTCGTCTGCAGCTGTAGTAGTTTACCATATCAAAATGTTAACTCA	300	
Db	311	AGTCAATTCGTGACTCCACTGTGGCNGTTGTGTTATGATATCAAAATGTTAACTCA	370	
Qy	301	TTCCAGCAAACTCAAAAGTCGATTGATGTGCAGAACAGAAAGAGAAAGTGATGTTATC	360	
Db	371	TTCCAGCAAACTCAAAAGTCGATTGATGTGCAGAACAGAAAGAGAAAGTGATGTTATC	430	
Qy	361	ATCACGCTAGTAGGAATACAAAGATCTTCTGTCACAAGGCGCAAGTGTCAAGTTGAGGAG	420	
Db	431	ATCATGCTAGTAGGAATAAACAAGATCTTCTGTCACAAGGCGCAAGTGTCAATTTGAGGAG	490	
Qy	421	GGAGAGAGAAAGCCAAAGCGCTGAATGTGTTATTCGTTTAACTAGGGCAAAACTGGA	480	
Db	491	GGAGAGAGAAAGCCAAAGAGCTGAATGTGTTATTCGTTTAACTAGGTGCAAAAGCTGGA	550	
Qy	481	TACAACTGAAGCAGCTCTTTCCAGCTGTAGCAGCAGCTTTGCCGGGAATGGAAGCAC	540	
Db	551	TACAACTGAAGCAGCTCTTTCCAGCTGTAGCAGCAGCTTTGCCGGGAATGGAAGCAC	610	
Qy	541	CAGGACCGAAGCAGAGAGAATCATGAGTGACATAAACTGGAAGAGCTTCAGGAGCAACA	600	

Db 611 CAGGACGAAGCAGAGAAGATATGATTGACATAAACTGGAAAGCCCTCAGGAGCAACCA 670
Qy 601 GTCAGCGAAGGGGTGTTCTCTGCTACTCTCCCATGTCAATCTTCAACCCCTTCCTCAGAAG 660
Db 671 GTCAGTGAAGGAGGCTGTTCTGCTGAATGTCTCCCTAGTCAATCTTCAA-CCTTCTTCAGAAG 729
Qy 661 CCCCCT 666
Db 730 CTCAC 735

RESULT 4

US-10-094-749-430
; Sequence 430, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KENICHI
; APPLICANT: IRIE, KYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 430
; LENGTH: 2456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-430

Query Match 52.6%; Score 401; DB 17; Length 2456;
Best Local Similarity 100.0%; Pred. No. 3.5e-115;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 361 ATCAGCTAGTAGGAATAGAACAGATCTTCTGACAGAGGCGAAGTCTCAGTTGAGGAG 420
Db 202 ATCAGCTAGTAGGAATAGAACAGATCTTCTGACAGAGGCGAAGTCTCAGTTGAGGAG 261
Qy 421 GGAGAGGAAGCAAGGGCTGAATGTTACGTTTATTGAAACTAGGGCAAAACTGGA 480
Db 262 GGAGAGGAAGCAAGGGCTGAATGTTACGTTTATTGAAACTAGGGCAAAACTGGA 321
Qy 481 TACAATGTAAGCAGCTCTTTTCGACGCTGACAGAGCTTTGCGGGGAATGGAAAGCACA 540
Db 322 TACAATGTAAGCAGCTCTTTTCGACGCTGACAGAGCTTTGCGGGGAATGGAAAGCACA 381
Qy 541 CAGACGGAAGCAGAGAAGACATGATGACATAAACTGGAAAGCCCTCAGGAGCAAAACA 600
Db 382 CAGACGGAAGCAGAGAAGACATGATGACATAAACTGGAAAGCCCTCAGGAGCAAAACA 441
Qy 601 GTCAGCGAAGGGGTGTTCTCTGCTACTCTCCCATGTCAATCTTCAACCCCTTCCTCAGAAG 660
Db 442 GTCAGCGAAGGGGTGTTCTCTGCTACTCTCCCATGTCAATCTTCAACCCCTTCCTCAGAAG 501

Qy 661 CCCCCTACTCTTCTTCAATGACTGCAGTGTGAATATTGGCTTGAACTTTCCCTTCATTATA 720
Db 502 CCCCCTTACTCTTCTTCAATGACTGCAGTGTGAATATTGGCTTGAACTTTCCCTTCATTATA 561
Qy 721 ATAAAGTTTTTGCAATTCATCATTTGCTGCCCTGTCTCGTGGAG 761
Db 562 ATAAAGTTTTTGCAATTCATCATTTGCTGCCCTGTCTCGTGGAG 602

RESULT 5

US-10-172-118-1656
; Sequence 1656, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1656
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 016577
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1656

Query Match 50.3%; Score 383.2; DB 17; Length 1266;
Best Local Similarity 75.6%; Pred. No. 1e-109;
Matches 475; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

Qy 1 ATGTCGCGGGCGGAGACTTCCGGAAATCCGCTCAGGAAATTCAGCTGGTGTTCTCTGGG 60
Db 338 ATGTCGCGAGGGGAGATTTGGGAATCCACTGAGAAATTCAGTTGGTGTCTTGGG 397
Qy 61 GAGCAAGCGTTGCAAGAATCTTTGATCACAGATTCAGGTATGACAGTTTGAACAAC 120
Db 398 GAGCAGAGCGTCGGGAAGACGCTCTCTGATTACGAGGTTCTATGACGACAGCTTCGACAAC 457
Qy 121 ACCTATCAGCAATAATTGGCATTTGATCTTTTATCAAAAACTATGTTGAGGAGTGA 180
Db 458 ACATACAGGCAACCAATTTGGGATTTGACTTCTGTCAAAAAACCATGTACTTTGGAGACCGC 517
Qy 181 ACAATCGGCTTCGGCTGTGGGATACGGCGGTCAGAAAGCTCCGTCAGCCTCATTTCCC 240
Db 518 ACGTTCGAGCTGACGCTCTGGGACAGCTGGTCAGAGAGGTTCCGACGCTGATCCCC 577
Qy 241 AGGTACATCCGTCGATTCTGCTGCAGCTGTAGTAGTTTACGATATCAAAATGTTAACTCA 300
Db 578 AGCTACATCCGGGACTCCACGGTGGCTGTGTGTGTACGACATCACAATCTCAACTCC 637
Qy 301 TTCCAGCAAACTCAAAAGTGGATTGATGTCTCAGAAACAGAAAGAGGAAGTGTATTC 360
Db 638 TTCCAACAGACCTCTAAGTGGATCGACGCTCAGGACAGAGAGGGGACGTGATGTATC 697
Qy 361 ATCAGCTAGTAGGAATAGACAGATCTTGTCTGACAAAGGCAAGTGTCAAGTTGAGGAG 420
Db 698 ATCATGCTGTGGGCAACAAAGCGGACCTGGCTGTATNAGAGGAGATTAACATCGAGAG 757
Qy 421 GGAGAGAGGAAGCAAGGGCTGAATGTTTACGTTTATTGAAACTAGGGCAAAACTGGA 480

Db 758 GGGAGCAGCGGCCAAAGAACTGAGCGTCATGTTTCATTGAGACCAGTCGGAAGACTGCG 817
QY 481 TACAATGTAAGCAGCTCTTTGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAGCACCA 540
Db 818 TACACGTGAGCAGCTTTTTCGACGTGTGGCTCTACCCGGAATGGAGATGTC 877
QY 541 CAGGACGGAAGCAGAGAACATGAGTGACATAAACTCGAAAAGCCTCAGGAGCAACA 600
Db 878 CAGGAGAAAAGCAAGAGGATGATCGACATCAAGCTGGACAAACCCCGAGGCCCCG 937
QY 601 CTCAGCGAAGGGGTGTTCTCTGCTACT 628
Db 938 GCCAGCGGGCGGCTGCTCTGCTAAT 965

RESULT 6

US-10-342-887-1656
; Sequence 1656, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.

; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699

; SEQ ID NO 1656
; LENGTH: 1266

; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-342-887-1656

Query Match 50.3%; Score 383.2; DB 17; Length 1266;
Best Local Similarity 75.6%; Pred. No. 1e-109;
Matches 475; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 1 ATGTCGCGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGGCTGGTGTCTCTGGGG 60
Db 338 ATGTCGCGAGGGGAGATTTTGGGAATCCACTGAGAAATTCAGTTGGTGTCTTGGGG 397
QY 61 GAGCAAAAGCGTTCGCAAGACATCTTTGATCACAAGATTCAGGTATGACAGTTTGCACAA 120
Db 398 GAGCAGAGCGTCGGGAAGACGTCCTGATTACAGGTTTCATGTACGACAGCTTCGACAA 457
QY 121 ACCTATCAGGCAATTAATGGCATTCGCTTTTATCAAAACTATGTTACTTGGAGGATGGA 180
Db 458 ACATACAGGCAACCAATTTGGGATTCGCTGATTACAGGTTTCATGTACGACAGCTTCGACAA 457
QY 121 ACCTATCAGGCAATTAATGGCATTCGCTTTTATCAAAACTATGTTACTTGGAGGATGGA 180
Db 458 ACATACAGGCAACCAATTTGGGATTCGCTGATTACAGGTTTCATGTACGACAGCTTCGACAA 517
QY 181 ACAATCGGCTTCGGCTGTGGGATACGGGGTTCAGGAACGTCCTCGTAGCTCTATCC 240
Db 518 AGGTGCGACTGCGGCTCTGGGACACAGCTGTGTCAGGAGGTTTCGCGAGCTGTATCC 577
QY 241 AGGTATACCTCGGATTCCTCGCAGCTGTAGTGTTCAGATATCAAAATGTTAACTCA 300
Db 578 AGCTATACCTCGGACTTCACCGTGGCTGTGGTGTGATACGACATCAAAATCTCAACTCC 637
QY 301 TTCCAGCAAACTACAAAGTGGATGATGTGTCAGAACAGAGGAGGAGTGTATTATC 360
Db 578 AGCTATACCTCGGACTTCACCGTGGCTGTGGTGTGATGACATCAAAATCTCAACTCC 637
QY 301 TTCCAGCAAACTACAAAGTGGATGATGTGTCAGAACAGAGGAGGAGTGTATTATC 360
Db 638 TTCCAAACAGACCTCTAAGTGGATTCGACGACGTCAGGACAGAGGGGCGAGTGTATTATC 697

QY 361 ATCAGCTAGTAGGAATAGAACAGATCTTGTCTGACAAGAGGCAAGTGTCACTTGAGGAG 420
Db 698 ATCATGTCTGTTGGGCAACAAGACGGACCTGGCTGATAAGAGGACAGATAACCATCGAGGAG 757
QY 421 GGAGAGAGAAAGCCAAAGGGCTGAATGTTAGCTTTTATTGAAACTTAGGGCAAAACTGGA 480
Db 758 GGGGAGCAGCGCGCCAAAGAACTGAGCGTCATGTTTCATTGAGACCAAGTCCGAAGACTGCG 817
QY 481 TACAATGTAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAGCACCA 540
Db 818 TACAACGTGAAGCAGCTTTTTCGACGTGTGGCTCGGCTCTACCCGGAATGGAGATGTC 877
QY 541 CAGGACGGAAGCAGAGAACATGAGTGACATAAACTGGAAGAGCCTCAGGAGCAACA 600
Db 878 CAGGAGAAAAGCAAGAGGATGATCGACATCAAGCTGGACAAACCCCGAGGCCCCG 937
QY 601 CTCAGCGAAGGGGTGTTCTCTGCTACT 628
Db 938 GCCAGCGGGCGGCTGCTCTGCTAAT 965

RESULT 7

US-10-848-755A-181

; Sequence 181, Application US/10848755A
; Publication No. US20050054826A1
; GENERAL INFORMATION:

; APPLICANT: Mao, Mao

; TITLE OF INVENTION: HUMAN DIAPHANOUS-3 GENE AND METHODS OF USE THEREFOR

; FILE REFERENCE: 9301-196-999

; CURRENT APPLICATION NUMBER: US/10/848,755A

; CURRENT FILING DATE: 2004-05-18

; PRIOR APPLICATION NUMBER: 60/471,842

; PRIOR FILING DATE: 2003-05-11

; NUMBER OF SEQ ID NOS: 275

; SOFTWARE: Patencin version 3.2 CAM: 301891-999188

; SEQ ID NO 181

; LENGTH: 1266

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-848-755A-181

Query Match 50.3%; Score 383.2; DB 19; Length 1266;
Best Local Similarity 75.6%; Pred. No. 1e-109;
Matches 475; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 1 ATGTCGCGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGGCTGGTGTCTCTGGGG 60
Db 338 ATGTCGCGAGGGGAGATTTTGGGAATCCACTGAGAAATTCAGTTGGTGTCTTGGGG 397
QY 61 GAGCAAAAGCGTTCGCAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTCGACAA 120
Db 398 GAGCAGAGCGTCGGGAAGACGTCCTGATTACGAGTTCATGTACGACAGCTTCGACAA 457
QY 121 ACCTATCAGGCAATTAATGGCATTCGCTTTTATCAAAACTATGTTACTTGGAGGATGGA 180
Db 458 ACATACAGGCAACCAATTTGGGATTCGCTTTGTCAAAACCAATGTACTTCTGGAGGACCG 517
QY 181 ACAATCGGCTTCGGCTGTGGGATACGGGGTTCAGGAACGTCCTCGTAGCTCTATCC 240
Db 518 AGGTGCGACTGCGGCTCTGGGACACAGCTGTGTCAGGAGGTTTCGCGAGCTGTATCC 577
QY 241 AGGTATACCTCGGATTCCTCGCAGCTGTAGTGTTCAGATATCAAAATGTTAACTCA 300
Db 578 AGCTATACCTCGGACTTCACCGTGGCTGTGGTGTGATACGACATCAAAATCTCAACTCC 637
QY 301 TTCCAGCAAACTACAAAGTGGATGATGTGTCAGAAACAGAAAGAGGAGTGTATTATC 360
Db 638 TTCCAAACAGACCTCTAAGTGGATTCGACGACGTCAGGACAGAGGGGCGAGTGTATTATC 697
QY 361 ATCAGCTAGTAGGAATAGAACAGATCTTGTCTGACAGAGGCAAGTGTCACTTGAGGAG 420
Db 698 ATCATGTCTGTTGGGCAACAAGACCGACCTGGCTGATTAAGAGGACAGATAACCATCGAGGAG 757

Qy	421	GGAGAGGAGNAAGCCAAAGGGCTGAATGTTATGCTTTATTGAAACTAGGGCAAAACCTGGA	480
Db	758	GGGGAGCAGCGGCCCAAAGAACTGAGCGTCATGTTTCATTGAGACAGTCGCGAAGACTGGC	817
Qy	481	TACAATGTTAAACGACGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAGCACA	540
Db	818	TACAACGTGAAGCAGCTTTTTCGACGTGTGGCGTCGGCTCTACCCGGAATGGAGAAATGTC	877
Qy	541	CAGGACGGAGACAGAGAAGACATGAGTCACATATAAACTGGAAGAACCTTCAGGAGCAACA	600
Db	878	CAGGAGAAAGCAAGAGGGATGATGCACATCAAGCTGGACAAACCCAGGAGGCCCCCG	937
Qy	601	GTACAGCAAGGGGGTTGTTCTCGCTACT	628
Db	938	GCCAGCGAGGGCGGCTGCTCTCTGCTAAT	965

```

RESULT 8
US-09-960-352-7355
; Sequence 7355, Application US/09960352
; Patent NO. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Machilagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7355
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 32-LIB188-014-Q1-E1-H11
US-09-960-352-7355

```

Query Match	50.1%;	Score 382;	DB 9;	Length 443;
Best Local Similarity	93.0%;	Pred. No. 1.3e-109;		
Matches 400;	Conservative 0;	Mismatches 30;	Indels 0;	Gaps 0;
QY	166	TACTTGGAGGATGCAACAATCGGGCTTCGGCTGTGGGATACGGCGGGTCAGGAACGATC	225	
Db	14	TACTTGGAAAGATCGAACAATCAGGCTGGAGCTGTGGGATCTCGGGTCAGGAACGTTTC	73	
QY	226	CGTAGCCTCATTTCCACAGTACATCCGTGATTTCTGCTGCAGCTGTAGTAGTTTACGATATC	288	
Db	74	CGTAGCCTCATTTCCAGTTACATCCGTGATTTCTGCTGCAGCTGTAGTAGTTTATGATATC	133	
QY	286	ACAAATGTTTAACCTATTTCCAGCAAACTACAAAATGGATTTGATGATGTGAGAACAGAAAGA	345	
Db	134	ACAAATGTTTAACCTCTTCCAGCAAACTACAAAATGGATTTGATGATGTGAGAACAGAAAGA	193	
QY	346	GGAGTGTATGTTATCATCACTAGTAGGAAATAGAACAGATCTTGCTGCAAGAGGCAA	405	
Db	194	GGAGTGTATGTTATCATCATGCTAGTAGGAAATAAAACAGATCTTGCTGCAAGAGGCAA	253	
QY	406	GTGTGAGTGTGAGGAGGAGAGGAAAGCCAAAGGGCTGAATGTTCAGTTTATTGAAACT	465	
Db	254	GTGTCAATCGAGGAAGAGAGAGGAAAGCCAAAGAGCTGAATGTTCAGTTTATTGAAACT	313	
QY	466	AGGCCAAAACCTGGATACAATGTAAAGAGCTCTTTTTCAGCTGTAGCAGAGCTTTGCCG	525	
Db	314	AGTGCAAAAGCAGGATACAATGTAAAGCAGCTCTTCCGACGTGTTCGACAGCTCTACCT	373	
QY	526	GGAAATGAAAGCACACAGGACGGAGAGAGACATGATGATACATAAACTCGAAAG	585	
Db	374	GGAAATGAAAGCACACAGACAGAGCAGAGACAGACATGATGATACATAAACTCGAAAAA	433	
QY	586	CCTCAGGAGC	595	

Db 434 CCTCAAGAGC 443

RESULT 9

US-10-017-161-1727

; Sequence 1727, Application US/10017161

; Publication No. US2003014369A1

; GENERAL INFORMATION:

; APPLICANT: SUWA, MAKIHO

; APPLICANT: ASAI, KIYOSHI

; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI

; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS

; FILE REFERENCE: 084335/0152

; CURRENT APPLICATION NUMBER: US/10/017,161

; CURRENT FILING DATE: 2002-12-18

; PRIOR APPLICATION NUMBER: JP 2001/246789

; PRIOR FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 2430

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1727

; LENGTH: 3826

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: source

; LOCATION: (1)..(3826)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (201)..(334)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (2878)..(2995)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (3081)..(3626)

; FEATURE:

; NAME/KEY: modified_base

; LOCATION: (3370)

; OTHER INFORMATION: a, t, c, g, unknown or other

; FEATURE:

; NAME/KEY: modified_base

; LOCATION: (3568)

; OTHER INFORMATION: a, t, c, g, unknown or other

; FEATURE:

; NAME/KEY: modified_base

; LOCATION: (3776)..(3826)

; OTHER INFORMATION: a, t, c, g, unknown or other

US-10-017-161-1727

Query Match	46.2%	Score 352.2;	DB 15;	Length 3826;
Best Local Similarity	93.1%;	Pred. No. 1.2e-99;		
Matches 444;	Conservative 0;	Mismatches 25;	Indels 8;	Gaps 7;
Qy	1	ATGTCCGCGGGCGGAGACTTTCGGGAATCCGCTCAGGAAATTCAGCTGTGTCTCTGGGG	60	
Db	3302	ATGTCCGCGGGC-GAGACTTTCGGGAATCCGCTCAGGAAATTCAGCTGTGTCTCTGGGG	3360	
Qy	61	GAGCAAGCGTTCGAAGACATCTTTGATCACAGATTTCAGGTATGACAGTGTGTGACAAC	120	
Db	3361	GAGCAAA-CTNTGCAAGACATCTTTGATCACAGATTTCAGGTATGACAGTGTGTGACAAC	3419	
Qy	121	ACCTATCAGGCAATAATGGCAATGTACTTTTATCAAAAACACTATGTACTTTGGAGGATGGA	180	
Db	3420	ACCTATCAGGCAATAATGGCAATGTACTTTTATCAAAAACATAATGATCTTTGGAGGATGGA	3479	
Qy	181	ACAATCGGGCTTCGGCTGTGGGATACGGCGGGTCAGGAACGCTCCGTAGACCTCATTTCCC	240	
Db	3480	ACAATCGGGCTTCGGCTGTGGGATACGGCGGGTCAGGAACGCTCCGTAGACCTCATTTCCC	3539	
Qy	241	AGGTACATCCGTGATTCTGCTGCAGCTGTAGTAGTTTACGATATCACAAAATGTTAACTCA	300	
Db	3540	AGGTACATCCGTGATTCTGCTGCAGCTGAGTAGTTTACGATATCACAAAATGTTAACTCA	3599	

QY 301 TTCCAGCAACTA-CAAACTGATTGA-TGATGTCAGAA-CAGAAAGAGGAAAGTGATGTT 357
D 3600 TTCCAGCAGACTACCAAAAGTGATTGACTGATGTCAGAACAGAGGAGGAGTGATGTT 3659
QY 358 ATCATCAGCTAGTAGGAATAAGAACAGATCTTGTCTGACAAAGAGGCAAGTGTCAGTTGAG 417
D 3660 ATCATCAGCTGTAGGCAA--TAAACAGATCTTGTCTGACAAAGAGGCAAGTGTCATTTGAG 3717
QY 418 GAGGAGAGAGGAAAGCCA-AGGGCTGAATGTTAGTTTATTGAAACTAGGGGCAA 473
D 3718 GCCGGAGAGCAGAAAGCCACAGGAGACATGTTACCTTTATTGAAACTGGGGCAA 3774

RESULT 10

US-10-292-798-1383
; Sequence 1383, Application US/10292798
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1383
; LENGTH: 3826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(3826)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(334)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2878)..(2995)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3081)..(3626)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (3370)..(3370)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (3568)..(3568)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (3776)..(3826)
; OTHER INFORMATION: a, t, c, g, unknown or other

US-10-292-798-1383
Query Match 46.2%; Score 352.2; DB 17; Length 3826;
Best Local Similarity 93.1%; Pred. No. 1.2e-99;
Matches 444; Conservative 0; Mismatches 25; Indels 8; Gaps 7;

QY 1 ATGTCGGGGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGATGTTCTCTGGGG 60
D 3302 ATGTCGGGGGGG-GAGACTTCGGGAATCCGCTGAGGAAATTCAGATGTTCTCTGGGG 3360
QY 61 GAGCAAGCGTTTGCAAGACATCTTTTGATCACCAGATTGAGTATGACAGTTTGTGACAAAC 120

D 3361 GAGCAAA-CTNTGCAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGTGACAAAC 3419
QY 121 ACCTATCAGGCAATAATTTGGCATTTGACATTTTATCAAAACATATCTACTTGGAGATGGA 180
D 3420 ACCTATCAGGCAATAATTTGGCATTTGACATTTTATCAAAACATATCTACTTGGAGATGGA 3479
QY 181 ACAATCGGGCTTCGGCTGTGGGATACGGGGGTGAGGAACGCTCTCCGTAGCCCTCAATTC 240
D 3480 ACAATCGGGCTTCGGCTGTGGGATACGGGGGTGAGGAACGCTCTCCGTAGCCCTCAATTC 3539
QY 241 AGGTATACCTCGTGAATTTGCTGCGAGCTGTAGTTTACGATATCAGAAATGTTAACTCA 300
D 3540 AGGTATACCTCGTGAATTTGCTGCGAGCTGTAGTTTACGATATCAGAAATGTTAACTCA 3599
QY 301 TTCCAGCAACTA-CAAAGTGGATTGA-TGATGTCAGAA-CAGAAAGAGGAAAGTGATGTT 357
D 3600 TTCCAGCAGACTACCAAAAGTGATTGACTGATGTCAGAACAGAGGAGGAGTGATGTT 3659
QY 358 ATCATCAGCTAGTAGGAATAAGAACAGATCTTGTCTGACAAAGAGGCAAGTGTCAGTTGAG 417
D 3660 ATCATCAGCTGTAGGCAA--TAAACAGATCTTGTCTGACAAAGAGGCAAGTGTCATTTGAG 3717
QY 418 GAGGAGAGAGGAAAGCCA-AGGGCTGAATGTTAGTTTATTGAAACTAGGGGCAA 473
D 3718 GCCGGAGAGCAGAAAGCCACAGGAGACATGTTACCTTTATTGAAACTGGGGCAA 3774

RESULT 11

US-09-918-995-5540
; Sequence 5540, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5540
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(424)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-5540

Query Match 44.9%; Score 341.8; DB 10; Length 424;
Best Local Similarity 89.8%; Pred. No. 6.3e-97;
Matches 380; Conservative 0; Mismatches 31; Indels 12; Gaps 1;

QY 334 AGAACAGAAAGAGGAGGAGTGATGTTATCATCAGCTAGTAGGAAATAGAACAGATCTTGT 393
D 1 AGAACAGAAAGAGGAGGAGTGATGTTATCANGCTAGTTTAAATATAAACAGATCTTGT 60
QY 394 GACAAAGGCGCAAGTGTCAGTTTGAGGAGGAGAGGAAAGCCAAAGGCTGAATGTTAGG 453
D 61 GACAAAGGCGCAAGTGTCAGTTTGAGGAGGAGAGGAAAGCCAAAGAGCTGAATGTTATG 120
QY 454 TTTATTGAACTAGGGCAAAACCTGGATACAAATGTAAAGAGAGCTCTTTCGAGCTAGCA 513
D 121 TTTATTGAACTAGTGCAAAAGCTGGATACAAATGTAAAGAGAGCTCTTTCGAGCTAGCA 180
QY 514 GCAGCTTTGCCGGGATCGGAAGCACAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 573
D 181 GCAGCTTTGCCGGGATCGGAAGCACAGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240

Db 861 CTG 863
 |||

Search completed: April 25, 2005, 13:07:29
Job time : 526 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 06:53:31 ; Search time 3096 Seconds
(without alignments)
9368.531 Million cell updates/sec

Title: US-09-441-857-11

Perfect score: 762

Sequence: 1 atgtccgcggcgaggactt.....tgctgcctgtctcgtggaga 762

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsa1:*

9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	761	99.9	3084	3 HSM801695	AL136727 Homo sapi
2	596.6	78.3	1928	3 AK084131	AK084131 Mus muscu
3	596.6	78.3	1955	3 AK051246	AK051246 Mus muscu
4	596.6	78.3	3404	3 AK083262	AK083262 Mus muscu
5	595	78.1	897	5 BQ422263	BQ422263 AGENCOURT
6	596	76.9	628	6 CB565915	CB565915 AGENCOURT
7	576.6	75.7	615	6 CB067426	CB067426 iq36f10.x
8	570.6	74.9	609	5 BQ632282	BQ632282 il25a01.x
9	566	74.3	790	5 B1694279	B1694279 603347724
10	565.4	74.2	599	7 CV025778	CV025778 3550 Full
11	544	71.4	656	1 A1435940	A1435940 th80c04.x
12	528	69.3	744	6 CB320435	CB320435 AGENCOURT
13	525.2	68.9	762	2 BE965054	BE965054 601658839
14	522.6	68.6	632	2 BF431157	BF431157 7008c05.x
15	512	67.2	792	6 CA316338	CA316338 UI-M-FW0
16	499	65.5	809	7 CN227000	CN227000 RJB008D12
17	492.4	64.6	637	4 B1602796	B1602796 603247154
18	490.2	64.3	554	1 A1660753	A1660753 we66h09.x
19	480.2	63.0	600	4 BG805486	BG805486 0693-37 M
20	470	61.7	871	5 BU381224	BU381224 603857076
21	469.2	61.6	880	4 B1602882	B1602882 603250358
22	469	61.5	769	6 CB988054	CB988054 AGENCOURT
23	467.4	61.3	648	7 CF172759	CF172759 B0912F09-
24	464.4	60.9	530	5 BQ086790	BQ086790 ih90b06.Y

25	464.4	60.9	534	5 BQ086839	BQ086839 ih90g05.Y
26	461.8	60.6	585	2 AW975757	AW975757 EST387866
27	461.6	60.6	518	4 BM833772	BM833772 K-EST0108
28	460.2	60.4	764	5 BU358285	BU358285 603476202
29	458.6	60.2	914	6 CA974052	CA974052 AGENCOURT
30	457.6	60.1	838	7 CK315856	CK315856 SB02028A1
31	457.4	60.0	572	7 CF198211	CF198211 maj45e04.
32	452.6	59.4	549	1 AL600766	AL600766 DKFZp313C
33	452	59.3	532	1 AA824586	AA824586 oc83c09.B
34	450.4	59.1	614	5 BX279398	BX279398 BX279398
35	449	58.9	563	7 CR556809	CR556809 DKFZp459E
36	449	58.9	804	7 CK302212	CK302212 SB02015A1
37	447.6	58.7	696	5 BX852687	BX852687 BX852687
38	446	58.5	883	4 BG108074	BG108074 602280008
39	445	58.4	655	8 AZ609020	AZ609020 IM0433M03
40	444.6	58.3	757	7 CK302742	CK302742 SB02016A2
41	444.4	58.3	878	6 CD361397	CD361397 AGENCOURT
42	441.2	57.9	858	7 CF547798	CF547798 AGENCOURT
43	439.4	57.7	759	5 BF709291	BF709291 BP709291
44	439.4	57.7	775	1 AI893257	AI893257 me39c01.Y
45	438.4	57.5	886	5 BU363505	BU363505 603785110

ALIGNMENTS

RESULT 1	HSM801695	3084 bp	mrna	linear	HTC 22-SEP-2004
LOCUS	Homo sapiens mRNA; cDNA DKFZp566K144 (from clone DKFZp566K144).				
DEFINITION	AL136727				
ACCESSION	AL136727.1	GI:12052972			
VERSION	HTC.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 3084)				
AUTHORS	Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.				
CONSRMT	The German cDNA Consortium				
TITLE	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764				
JOURNAL	Neuherberg, GERMANY				
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.				
	This clone (DKFZp566K144) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp566K144 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.				
FEATURES	Location/Qualifiers				
source	1. 3084				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="RZPD:DKFZp566K144"				
	/db_xref="taxon:9606"				
	/clone="DKFZp566K144"				
	/tissue type="kidney"				
	/clone_lib="566 (synonym: hfk22). Vector pAMP1; host Xl-2blue; sites NotI + SalI"				
	/dev stage="fetal"				
	/note="RAB6C, member RAS oncogene family"				
gene	1. 3084				
	/gene="DKFZp566K144"				
	456..1220				
	/gene="DKFZp566K144"				
CDS	/codon_start=1				
	/product="hypothetical protein"				

```
/protein_id="CAB66661.1"
/db_xref="GI:12052973"
/db_xref="GOA:Q9H0N0"
/translation="MSAGDFGNPLRKFVLFLGQSVAKTSLITFRFYDSFDNTYQA
IIGDFLSKTYLEDGTIGLRWDTAGORLSLIPYIRDSAAAAYVYDITVNSFO
OTTWIDVRTERGSDVLIITLVGNRTDLADKROYSVEGERKAGLNVTFLETRAKTG
YNNQLFRVAAALPCMSTQDGRMSDIKLEKPBQETVSEGGSCSYSPSSSTLP
QKPPYSFIDCSVNIGLNLPFLSLITFCNSLLLPVSMR"
```

ORIGIN	Query Match		99.9%; Score 761; DB 3; Length 3084;			
	Best	Local	Similarity	100.0%;	Pred. No.	5.6e-217;
	Matches	761;	Conservative	0;	Mismatches	0;
					Indels	0; Gaps
QY	1	ATGTCGGGGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGACTGGTGTCTCTGGGG	60			
Db	456	ATGTCGGGGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGACTGGTGTCTCTGGGG	515			
QY	61	GAGCAAGCGTTGCAAGACATCTTTGATCACAGATTGAGTATGACAGATTTTGACAAAC	120			
Db	516	GAGCAAGCGTTGCAAGACATCTTTGATCACAGATTGAGTATGACAGATTTTGACAAAC	575			
QY	121	ACCTATCAGGCAATAATTGGCATTCACCTTTTATCAAAACTATGTACTTCGAGGATGGA	180			
Db	576	ACCTATCAGGCAATAATTGGCATTCACCTTTTATCAAAACTATGTACTTCGAGGATGGA	635			
QY	181	ACAATCGGGCTTCGGCTGTGGGATACGGCGGCTCAGGAACGTCCTCGTAGCTCATTTCCC	240			
Db	636	ACAATCGGGCTTCGGCTGTGGGATACGGCGGCTCAGGAACGTCCTCGTAGCTCATTTCCC	695			
QY	241	AGGTATATCCGTGATCTCTGTCAGCTGTAGTGTAGTTTACGATATCACAAATGTTAACTCA	300			
Db	696	AGGTATATCCGTGATCTCTGTCAGCTGTAGTGTAGTTTACGATATCACAAATGTTAACTCA	755			
QY	301	TTCCAGCAAACTACAAAGTGATTCATGATGTGCAACAGAAAGGAGGTGATGTTATC	360			
Db	756	TTCCAGCAAACTACAAAGTGATTCATGATGTGCAACAGAAAGGAGGTGATGTTATC	815			
QY	361	ATCAGCTAGTAGGAAATAGAACAGATCTTCTGCAAGAGGCAAGTGTCAAGTGTGAGGAG	420			
Db	816	ATCAGCTAGTAGGAAATAGAACAGATCTTCTGCAAGAGGCAAGTGTCAAGTGTGAGGAG	875			
QY	421	GGAGAGGAAAGCAAGGGCTGAATGTTAGCTTTATTTGAACATAGGGCAAAACTGGA	480			
Db	876	GGAGAGGAAAGCAAGGGCTGAATGTTAGCTTTATTTGAACATAGGGCAAAACTGGA	935			
QY	481	TACAAATGTAAGCAGCTCTTTTCGAGCTGTAGCAGCAGCTTTTCCGGGAATGGAAGCACA	540			
Db	936	TACAAATGTAAGCAGCTCTTTTCGAGCTGTAGCAGCAGCTTTTCCGGGAATGGAAGCACA	995			
QY	541	CAGGACGGAAGCAGAGAGACATGATGACATAAACTGGAAAACGCTCAGGAGCAACA	600			
Db	996	CAGGACGGAAGCAGAGAGACATGATGACATAAACTGGAAAACGCTCAGGAGCAACA	1055			
QY	601	GTCAGGAGGGGGTGTCTCTGCTACTCTCCCATGTCTTCAACCCCTTCCTCAGAG	660			
Db	1056	GTCAGGAGGGGGTGTCTCTGCTACTCTCCCATGTCTTCAACCCCTTCCTCAGAG	1115			
QY	661	CCCCCTTACTCTTTTCATTCAGTGCAGTGTGAATATTGGCTTTGAACCTTTCCCTTCATTA	720			
Db	1116	CCCCCTTACTCTTTTCATTCAGTGCAGTGTGAATATTGGCTTTGAACCTTTCCCTTCATTA	1175			
QY	721	ATAAGCTTTTGCATTCATCATTTGCTGCTGTCTCGTGAG	761			
Db	1176	ATAAGCTTTTGCATTCATCATTTGCTGCTGTCTCGTGAG	1216			

RESULT 2
AK084131
LOCUS
DEFINITION Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:DJ30097N11 product:RAB6, member RAS

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

oncogene family, full insert sequence.
AK084131
GI:26350968
HTC; CAP trapper.
Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata,K., Itoh,M., Aizawa,K., Nagao,K., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1928)

Fukuda,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES

source

Location/Qualifiers

1..1955
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM DB:D130023A03"
/db_xref="taxon:10090"
/clone="D130023A03"
/tissue_type="spinal ganglion"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="12 days embryo"
474..1100
/note="unnamed protein product: RAB6, member RAS oncogene family [MGP|MGI:894313, GB|NM_024287, evidence: BLASTN, 99%, match=1919]
putative"
/codon_start=1
/protein_id="BAC34572.1"
/db_xref="GI:26341820"
/translation="MSAGGDSGNFLKFLVPLGDSVGKTSILTRFMYSFNDTYQA
QTIDFSLKWLSDRLVRLQMDKAGDERFSLIPSYIRDSVAVVYDITNNSFQ
TQTIDWDRTERSDVIMLVNKTDLADKQVSIIEGERKAKLVNMFYIETSAKAG
YNVQLFRVAAALPGMESTQDRSDRMIDIKLEKPOEPVNEGCS"

CDS

1..1955
/note="unnamed protein product: RAB6, member RAS oncogene family [MGP|MGI:894313, GB|NM_024287, evidence: BLASTN, 99%, match=1919]
putative"
/codon_start=1
/protein_id="BAC34572.1"
/db_xref="GI:26341820"
/translation="MSAGGDSGNFLKFLVPLGDSVGKTSILTRFMYSFNDTYQA
QTIDFSLKWLSDRLVRLQMDKAGDERFSLIPSYIRDSVAVVYDITNNSFQ
TQTIDWDRTERSDVIMLVNKTDLADKQVSIIEGERKAKLVNMFYIETSAKAG
YNVQLFRVAAALPGMESTQDRSDRMIDIKLEKPOEPVNEGCS"

ORIGIN

Query Match 78.3%; Score 596.6; DB 3; Length 1955;
Best Local Similarity 86.9%; Pred. No. 1.3e-167;
Matches 672; Conservative 0; Mismatches 89; Indels 12; Gaps 1;
QY 1 ATGTCGGCGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTCTCTCTGGG 60
DB 474 ATGTCGGCGGGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTCTCTCTGGG 533
QY 61 GAGCAAGCGTTGCAAGACATCTTTGATACACAGATTCAGGTATGACAGTTTTCAGAAC 120
DB 534 GAGCAGAGCGTTGGAAGACGCTCTTGATCACCCGATTCATGTATGACAGTTTTCAGAAC 593
QY 121 ACCTATCAGGCAATATTCGATTCGATCTTTTATCAAAATCTGCTGAGGATGGA 180
DB 594 ACCTATCAGGCAATATTCGATTCGATCTTTTATCAAAATCTGCTGAGGATGGA 653
QY 181 ACAATCGGGCTTCGGCTGCGGATACGGCGGTGAGGAACGCTCCGCTGAGCTCATTCCT 240
DB 654 ACCGTGCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCCT 713
QY 241 AGGTATACCGGTGATTCGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATCA 300
DB 714 AGCTACATTCGAGACTCCACTGTGCGAGTCTGTTGTTATGATATACAAATGTTTAACTCA 773
QY 301 TTCAGCAAACTACAAAGTGATTCATGATCTCAGACAGAAAGAGGAGTGATGTTATC 360
DB 774 TTCAGCAAACTACAAAGTGATTCATGATCTCAGACAGAAAGAGGAGTGATGTTATC 833
QY 361 ATCAGCTAGTAGGAAATAGAACAGATCTTCTGTCAGAGAGGCAAGTGTCTAGTTGAGGAG 420

Db 834 ATCATGCTAGTAGGAAATATAAACACAGATCTTCTGATAGAGGCAAGTGTCTCAATTGAGGAG 893
QY 421 GGAGAGAGGAAAGCCAAAGGGCTGAATGTTACGTTTATTTGAAACTAGGGCAAACTGGA 480
Db 894 GGAGAGAGGAAAGCCAAAGAGCTGAATGTTATGTTTATTTGAAACAGTGTGCAAAAGCAGGA 953
QY 481 TACAATCTAAAGCAGCTCTTTTCGACGCTGTAGCAGCAGCTTTCCGCGGAATCGGAAGCACA 540
Db 954 TACAATCTAAAGCAGCTCTTTTCGCGCGCTGTCTGCTGAGCTTTTACCTGGAATGGAAGCACA 1013
QY 541 CAGGACGAGAGCAGAGAGACATGAGTGACATAAACTGGAAAAAGCCTCAGGAGCAACA 600
Db 1014 CAGGACGAGAGCAGAGAGACATGAGTGACATAAACTGGAAAAAGCCTCAGGAGCAACA 1073
QY 601 GTACAGCAGAGGGGTTGTTCTGCTACTCTCCCATGTCATCTTCAACCCCTTCCTCAGAG 660
Db 1074 GTCAATGAAGGCGCTGTCTGCTGCTAAATCCCTCGGCATCTTTTCAAGCCCTCTCGAGAG 1133
QY 661 C-----CCCCCTTACTCTTTCATTCGACGCTGGAATATTTGGCTTCAACCTT 708
Db 1134 CTCACTGCTTTGGCCCCCATCTCTTTCATGACTGCGAGTGGAATATTTGGCTTGAACCTT 1193
QY 709 TTCCCTTCATTAATAACGTTTTCGCAATTCATTCATTCGCTGCTGCTGCTGAG 761
Db 1194 TTCCCTTCAGTAATAACGTTTTCGCAATTCATTCATTCGCTGCTGCTGCTGAG 1246
RESULT 4
AK083262 3404 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus adult male hippocampus cDNA, RIKEN full-length
DEFINITION enriched library, clone: C630031P05 product: RAB6, member RAS
oncogene family, full insert sequence.
ACCESSION AK083262
VERSION AK083262.1 GI:26350388
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
93279253
10349636
REFERENCE 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
REFERENCE 4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	
	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
	Nature 420, 563-573 (2002)	
	6 (bases 1 to 3404)	
REFERENCE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.	
	Direct Submission	
	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	
COMMENT	Please visit our web site for further details.	
	URL: http://genome.gsc.riken.jp/	
	URL: http://fantom.gsc.riken.jp/	
	Location/Qualifiers	
FEATURES	source	
	1..3404	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
CDS	/strain="C57BL/6J"	
	/db_xref="PANTOM DB:C630031P05"	
	/db_xref="taxon:10090"	
	/clone="C630031P05"	
polyA_signal	/sex="male"	
	/tissue_type="hippocampus"	
	/note="unnamed protein product; RAB6, member RAS oncogene family (MGD MGI:894313, GB NM_024287, evidence: BLASTN, 99%, match=1919)"	
	/dev_stage="adult"	
polyA_site	453..1089	
	/note="unannotated protein product; RAB6, member RAS oncogene family (MGD MGI:894313, GB NM_024287, evidence: BLASTN, 99%, match=1919)"	
	/codon_start=1	
	/protein_id="BAC38834.1"	
ORIGIN	/db_xref="GI:26350389"	
	/translation="MSAGDGFNGLRFLKFLVGLGQSVGKTSLSITRFMYDSFDNTYQA TIGDLSKTYMLEDTRVLQDLTAGQERFLPSYIRDSITVAVVVDITVNSFQ OTTKWIDVTRGSDVIIMLVGNKTDLDKQVSIERGERKAKELNWFETSAKAG YNNKQLEFRVAALPGMESTQDRSDRIDIKLEKPPQVNEGSCS"	
	3379..3384	
	/note="putative"	
Query Match	3404	
	/note="putative"	
	Best Local Similarity 86.9%; Pred. No. 1.6e-167;	
	Matches 672; Conservative 0; Mismatches 89; Indels 12; Gaps 1;	
Db	1 ATGTCGCGGGCGGAGCTTCGGGAATCCGCTGAGGAATTCAGGTGCTGCTCTGGGG 60	
	463 ATGTCGCGGGCGGAGCTTCGGGAATCCGCTGAGGAATTCAGGTGCTGCTCTGGGA 522	
Qy	61 GAGCAAGAGCTTGCAGGATCTTTGATCACCAGATTCAGGTATGACAGTTCGACAC 120	
	523 GAGCAGAGCGTTGGAAAGACGCTCTTGATCACCAGATTCATGTATGACAGTTCGACAC 582	
Qy	121 ACCATCAGGCAATTAATTTGGCATTGACATTTTATCAAAAACATGATGCTGGAGATGA 180	
	583 ACCATCAGGCAATTAATTTGGCATTGACATTTTATCAAAAACATGATGCTGGAGATGA 642	
Qy	181 ACAATCGGGCTTCGGCTCTGGGATACGGCGGTGAGGAACGCTCTCCGTAGCTCATTTCCC 240	
	643 ACCGTGCAATTCGAATTTAGGACACAGAGGTCAAGAGCGGTCAGGAGCTTGATTCCT 702	
Qy	241 AGGTACATCCGTGATTTCTGCTGAGCTGTAGTGTATGATATACAAAATGTTAACTCA 300	
	703 AGCTACATTCGAGACTCCACTGTGCGAGTTGTTGTTATGATATCACAATGTTAACTCA 762	
Qy	301 TTCCAGCAAACTACAAAGTGGATTCATGATCTCAGAACAGAGGAGGAGTGTATGTC 360	
	763 TTCCAGCAAACTACAAAGTGGATTCATGATCTCAGAACAGAGGAGGAGTGTATGTC 822	
Qy	361 ATCAGCTAGTAGGAATAGAACAGATCTTCTGCTCAGAACAGGCAAGTGTCTGAGGAG 420	
	823 ATCATGCTAGTAGGAATAGAACAGATCTTCTGCTCAGAACAGGCAAGTGTCTGAGGAG 882	
Qy	421 GGAGAGAGAAAGCCAAAGGCTGAATGTTACGTTTATTGAACTAGGCGCAAACTGGA 480	
	883 GGAGAGAGAAAGCCAAAGGCTGAATGTTACGTTTATTGAACTAGGCGCAAACTGGA 942	
Qy	481 TACATGTTAAAGCAGCTCTTTTCGAGCTGTACGACAGCTTTCCGCGGATGGAAGCACA 540	
	943 TACATGTTAAAGCAGCTCTTTTCGAGCTGTACGACAGCTTTCCGCGGATGGAAGCACA 1002	
Qy	541 CAGGACGGAAGCAGAGAGACATGATGACATATAAACTGGAAGAGCCTCAGGAGCAACA 600	
	1003 CAGGACGGAAGCAGAGAGACATGATGACATATAAACTGGAAGAGCCTCAGGAGCAACA 1062	
Qy	601 GTCAGGAGAGGGGTTGTTCTGCTGCTCTCCCATGTCATCTTCAACCTCTCTCAGAG 660	
	1063 GTCAATGAAGCGGCTGTTCTCTGCTTAATCCCTCGCATCTTTCAACCTCTCTCAGAG 1122	
Qy	661 C-----CCCCCTTACTCTTCAATTCACGTCAGTGTGAATATTTGGCTTGAACCTT 708	
	1123 CTCAGTCTGTTGGGCCCCATCTCTTCAATTCAGTGTGAATATTTGGCTTGAACCTT 1182	
Qy	709 TTCCCTTCATTAATAACGTTTTCGAAATTCATTCATTCGCTCTGCTCTGAGGAG 761	
	1193 TTCCCTTCATTAATAACGTTTTCGAAATTCATTCATTCGCTCTGCTCTGAGGAG 1235	
RESULT 5	BO422263	
	AGENCOURT_7802212 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6042084	
	5', mRNA sequence.	
	BO422263	
*ACCESSION	BO422263.1	
	GI:21117590	
	EST.	
	Homo sapiens (human)	
SOURCE	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 897)	
	NIH-MGC http://mgi.nci.nih.gov/.	
REFERENCE	National Institutes of Health, Mammalian Gene Collection (MGC)	
	Unpublished (1999)	
	Contact: Robert Strausberg, Ph.D.	
	Email: cgapbs-remail.nih.gov	
JOURNAL	Tissue Procurement: ATCC	
	cDNA Library Preparation: Life Technologies, Inc.	
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
	DNA Sequencing by: Agencourt Bioscience Corporation	
COMMENT	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:	
	http://image.llnl.gov	

Qy	301	TTCCGCAAACTACAAAGTGGATTGATGTCAGAACAGAGAGGAAGTGATGTATC	360
Db	301	TTCCGCAAACTACAAAGTGGATTGATGTCAGAACAGAGAGGAAGTGATGTATC	360
Qy	361	ATCAGCTAGTAGGAAATAGAACAGATCTTGCTGCACAAGAGGCAAGTGTCAAGTGGAGG	420
Db	361	ATCATGCTAGTAGGAAATAGAACAGATCTTGCTGCACAAGAGGCAAGTGTCAA	420
Qy	421	GGAGAGAGGAAGCCAAAGGGCTGAAGTGTACGTTTATTCGAACTAGGCGAAAAC	480
Db	421	GGAGAGAGGAAGCCAAAGAGCTGAATGTTATGTTTATTTGAAACTAGTGC	480
Qy	481	TACAAATGTAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAAT	540
Db	481	TACAAATGTAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAAT	540
Qy	541	CAGGACGGAAGCAGAGAGACATCAGTGACATAAACTGGAAAGCCTCAGGAGCA	600
Db	541	CAGGACAGAGCAGAGAGATATGATTGCATAAACTGGAAAGCCTCAGGAGCA	600
Qy	601	GTCCAGCAAGGGGGTTGTTCTCGCTA	626
Db	601	GTCCAGTGAAGGAGGCTGTTCTCGCTA	626

RESULT 7	CB067426	615 bp	mRNA	linear	EST 21-JAN-2003
LOCUS					
DEFINITION	i336f10.x1 HR85 islet Homo sapiens CDNA clone IMAGE: 3', similar to SW:RAB6 HUMAN P20340 RAS-RELATED PROTEIN RAB-6. [1] ;, mRNA sequence.				

FEATURES
SOURCE

RESULT 8
B0632282

FOCUS DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permut Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@ingate.wustl.edu, Tel: 314-363-1916, Fax: 314-747-3692."

ORIGIN

Query Match	75.7%	Score 576.6	DB 6	Length 615
Best Local Similarity	96.1%	Pred. No. 9,7e-162		
Matches 591	Conservative	0	Mismatches 24	Indels 0
Gaps	0			
Qy	10	GGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAAAGCTGGTGTTCCTGGGGGAGCAAAAGC	69	
Db	1	GGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAAAGCTGGTGTTCCTGGGGGAGCAAAAGC	60	
Qy	70	GTTTGAAAGACATCTTTTGATCACAGATTCAGGTATGACAGTTTGGACAAACCTATCAG	129	
Db	61	GTTTGAAAGACATCTTTTGATCACAGATTCAGGTATGATGACAGTTTGGACAAACCTATCAG	120	
Qy	130	GCAATAATTGGCAATTCGACTTTTATCAAAACATCTGTACTTGGAGGATGAAACAATCGGG	189	
Db	121	GCAACAATTGGCAATTCGACTTTTATCAAAACATCTGTACTTGGAGGATCGAACAATCAGG	180	
Qy	190	CTTCGGCTGTGGGATACGGCGGTGAGGAACTCTCCGTAGCTCATTTCCAGGTACATC	249	
Db	181	CTTCAGCTGTGGGATACTGCGGTGAGGAACTCTCCGTAGCTCATTTCCAGTTACATC	240	
Qy	250	CGTGATCTTCGCTGACGCTGTAGTGTTCAGTATCACAAATGTTAACTCATTTCCAGCAA	309	
Db	241	CGTGATCTTCGCTGACGCTGTAGTGTTCAGTATCACAAATGTTAACTCATTTCCAGCAA	300	
Qy	310	ACTACAAAGTGGATTGATGATGTGAGAACAGAAAGAGGAAGTGATGTATCATCAGCTA	369	
Db	301	ACTACAAAGTGGATTGATGATGTGAGAACAGAAAGAGGAAGTGATGTATCATCTGCTA	360	
Qy	370	GTAGGAAATAGAACAGATCTTGTGTGACAGAGCAAGTGTCAAGTGGAGGGAGAGAGG	429	
Db	361	GTAGGAAATAAACAGATCTTGTGTGACAGAGCAAGTGTCAATGTAGGAGGGAGAGAGG	420	
Qy	430	AAAGCCAAAGGGCTGAATGTTAGCTTTATTGAAACTTAGGGCAAAAACCTGGATACAATGTA	489	
Db	421	AAAGCCAAAGAGCTGAATGTTATGTTTATTGAAACTTAGTGCAAAAGCTGGATACAATGTA	480	
Qy	490	AAGCAGCTCTTTCGACGCTGTAGCAGCAGCTTTTCCCGGGAATGGAAGACACAGGACGGA	549	
Db	481	AAGCAGCTCTTTCGACGCTGTAGCAGCAGCTTTTCCCGGGAATGGAAGACACAGGACAGA	540	
Qy	550	AGCAGAGAGACATGATGATGACATAAAACCTGGAAGAGCCTCAGGAGCAAAACAGTCCAGCGAA	609	
Db	541	AGCAGAGAGATATGATTGACATAAAACCTGGAAGAGCCTCAGGAGCAAAACAGTCCAGTGAA	600	
Qy	610	GGGGGTGTTCTTCG	624	
Db	601	AGAGGCTGTTCTTCG	615	

BQ632282 609 bp mRNA linear EST 02-JUL-2002
i125a01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6030984 3'
similar to SW:RAB6_HUMAN P20340 RAS-RELATED PROTEIN RAB-6. [1] ;
mRNA sequence.

mRNA sequence.
 BQ632282
 BQ632282.1 GI:21683800
 EST.

Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Williams, T., Jackson, Y., McCann, R., Cole, R., Tsagarishvili, R., Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Possible reversed clone: similarity on wrong strand
Seq primer: -400P from Gibco
High quality sequence stop: 439.
Location/Qualifiers
1..609
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6030984"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

16 GACTTCGGGAATCCGCTCAGGAAATTCGAAGCTGGTTCCTGGGGGAGCAAGCGTTGCA 75
1 GACTTCGGGAATCCGCTCAGGAAATTCGAAGCTGGTTCCTGGGGGAGCAAGCGTTGGA 60
76 AAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGTGACACACCTATCAGGCAATA 135
61 AAGACATCTTTGATCACCAGATTGATGATGACAGTTTGTGACACACCTATCAGGCAACA 120
136 ATTGCGATTGCTTTTATCAAAAACATGATGACTTGGAGGATGGAACAAATCGGCTTCGG 195
121 ATTGCGATTGCTTTTATCAAAAACATGATGACTTGGAGGATGGAACAAATCAGGCTTCAG 180
196 CTGTGGGATACGGCGGGTCAGGAACGCTCCGCTAGCCCTCATTTCCAGGTACATCCGTGAT 255
181 CTGTGGGATACGGCGGGTCAGGAACGCTTCGCTAGCCCTCATTTCCAGTTACATCCGTGAT 240
256 TCTGCTGAGCTGTAGTAGTTTACGATATACAAATGTTAACTATTCACCAACATACA 315
241 TCTGCTGAGCTGTAGTAGTTTACGATATACAAATGTTAACTATTCACCAACATACA 300
316 AAGTGGATTGATGTGTCAGAACAGAAAGAGAGTGTATGTTATCATCACCTAGTAGGA 375
301 AAGTGGATTGATGTGTCAGAACAGAAAGAGAGTGTATGTTATCATCATCTAGTAGGA 360
376 AATGAACAGATCTTTGCTGACAAAGGCAAGTGTTCAGTTGAGGGGAGAGAGAAAGCC 435
361 AATAAAACAGATCTTTGCTGACAAAGGCAAGTGTCAATTGAGGGGAGAGAGAAAGCC 420

ORIGIN
Query Match 74.9%; Score 570.6; DB 5; Length 609;
Best Local Similarity 96.1%; Pred No. 6.2e-160;
Matches 585; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 16 GACTTCGGGAATCCGCTCAGGAAATTCGAAGCTGGTTCCTGGGGGAGCAAGCGTTGCA 75
Db 1 GACTTCGGGAATCCGCTCAGGAAATTCGAAGCTGGTTCCTGGGGGAGCAAGCGTTGGA 60
QY 76 AAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGTGACACACCTATCAGGCAATA 135
Db 61 AAGACATCTTTGATCACCAGATTGATGATGACAGTTTGTGACACACCTATCAGGCAACA 120
QY 136 ATTGCGATTGCTTTTATCAAAAACATGATGACTTGGAGGATGGAACAAATCGGCTTCGG 195
Db 121 ATTGCGATTGCTTTTATCAAAAACATGATGACTTGGAGGATGGAACAAATCAGGCTTCAG 180
QY 196 CTGTGGGATACGGCGGGTCAGGAACGCTCCGCTAGCCCTCATTTCCAGGTACATCCGTGAT 255
Db 181 CTGTGGGATACGGCGGGTCAGGAACGCTTCGCTAGCCCTCATTTCCAGTTACATCCGTGAT 240
QY 256 TCTGCTGAGCTGTAGTAGTTTACGATATACAAATGTTAACTATTCACCAACATACA 315
Db 241 TCTGCTGAGCTGTAGTAGTTTACGATATACAAATGTTAACTATTCACCAACATACA 300
QY 316 AAGTGGATTGATGTGTCAGAACAGAAAGAGAGTGTATGTTATCATCACCTAGTAGGA 375
Db 301 AAGTGGATTGATGTGTCAGAACAGAAAGAGAGTGTATGTTATCATCATCTAGTAGGA 360
QY 376 AATGAACAGATCTTTGCTGACAAAGGCAAGTGTTCAGTTGAGGGGAGAGAGAAAGCC 435
Db 361 AATAAAACAGATCTTTGCTGACAAAGGCAAGTGTCAATTGAGGGGAGAGAGAAAGCC 420

QY 436 AAAGGGCTGAATGTTACGTTTATTGAAACTAGGCGCAAAAACCTGGATACAAATGTTAAAGCAG 495
Db 421 AAAGGGCTGAATGTTACGTTTATTGAAACTAGTGCAAAAGCTGGATACAAATGTTAAAGCAG 480
QY 496 CTCCTTTCCAGCTGTAGCAGCAGCTTTTCGCCGGGAATGAAAGCACACAGGACGGAAGCAGA 555
Db 481 CTCCTTTCCAGCTGTAGCAGCAGCTTTTCGCCGGGAATGAAAGCACACAGGACGGAAGCAGA 540
QY 556 GAAGACATGAGTGACATAAACTGGAAGAGCTCAGGAGCAACAGTCAGGCAAGGGGGT 615
Db 541 GAAGATATGATTGACATAAACTGGAAGAGCTCAGGAGCAACACCTAGTCAGTGAAGAGGC 600
QY 616 TGTTCCTGTC 624
Db 601 TGTTCCTGTC 609

RESULT 9
BI694279
LOCUS
DEFINITION
603347724F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5375477 5', mRNA sequence.
ACCESSION
BI694279
VERSION
BI694279.1 GI:15656908
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 790)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11955 row: p column: 06
High quality sequence stop: 788.
Location/Qualifiers
1..790
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5375477"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam2"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN
Query Match 74.3%; Score 566; DB 4; Length 790;
Best Local Similarity 85.9%; Pred. No. 1.6e-158;
Matches 656; Conservative 0; Mismatches 95; Indels 13; Gaps 2;

QY 11 GCGGAGACTTCGGGAATCCGCTCAGGAAATTCAGCTGGT-GTTCCTGGGGAGCAAGC 69
Db 1 GCGGAGACTTCGGGAATCCGCTCAGGAAATTCAGCTGGTGGTTCCTGGGAGCAGCAGC 60
QY 70 GTTGCAAGACATCTTTTGTATGATCAACAGATTGAGTATGACAGTTTGTGACACCTATCAG 129
Db 61 GTTGCAAGACATCTTTTGTATGATCAACAGATTGATGACAGTTTGTGACACCTATCAG 120

```
QY 130 GCAATAATGGCATTTGACCTTTTATCAAAACATATGTACTTGGAGATGGAACTCGG 189
Db |||||
QY 180 CTTCCGGCTGTGGATACGGCGGTGAGAAAGTCTCCGCTCATTTCCAGGTACATC 249
Db |||||
QY 250 CGTGATTCGTGACGTGTAGTGTACGATATACAAATGTTAACTCATTTCCAGCA 309
Db |||||
QY 310 ACTCAAAAGTGGATGTGATGTGATGATGATGATGATGATGATGATGATGATGAT 369
Db |||||
QY 370 GTAGAAATAGAAAGATCTTCTGATGATGATGATGATGATGATGATGATGATGAT 429
Db |||||
QY 430 AAAGCAAGGCTGAATGTTAGCTTTTATGAACTAGGGCAAACTGGATACATGTA 489
Db |||||
QY 490 AAGCAGCTCTTTCGACGTGTAGCAGCGCTTTCCGGGAATGGAAGCACACAGCAGCA 549
Db |||||
QY 550 AGCAGAGAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
Db |||||
QY 610 GGGGTTGTTCTGCTACTCTCCCATGCTTCAACCTTCTCAGAGC----- 661
Db |||||
QY 662 ---CCCTTACTCTTTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 717
Db |||||
QY 718 TTAATAAGCTTTTGAATTCATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 761
Db |||||
```

```
RESULT 10
CV025778
LOCUS 3550 Full Length cDNA from the Mammalian Gene Collection Homo sapiens cDNA 5' similar to BC003617, mRNA sequence.
DEFINITION CV025778
ACCESSION CV025778.1 GI:51483727
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S., Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O., Clingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T., Simmonds,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C., Vandenhaute,J., Cusick,M.E., Albaladejo,J.S., Hill,D.E. and Vidal,M.
Human ORFome Version 1.1: a Platform for Reverse Proteomics
Genome Res. (2004) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc_Vidal@dfci.harvard.edu
```

ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF results from a PCR reaction using an MGC full-length cDNA as template DNA and ORF specific primers

PCR Primers

FORWARD: ATGTCCACGGCGGAGACTT
BACKWARD: TAGCAGAAACAGCTCTT

Insert Length: 599 Std Error: 42.00
Plate: 11029 row: 01 column: D

Seq primer: ACTGCGTCTGTTTACAACTGCTGACTGGGAAAAAC
High quality sequence start: 97
High quality sequence stop: 598
POLYA=No.

FEATURES

source

Location/Qualifiers

1..599
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="mixed"
/clone_lib="Full Length cDNA from the Mammalian Gene Collection"
/note="Vector: mixed; The ORFs were PCR amplified from the MGC (Mammalian Gene Collection) as of April 2004 and cloned by recombinational gateway cloning into pDONR223 Donor vector. Reference : MGC (Mammalian Gene Collection) Program Team, Generation and Initial Analysis of more than 15,000 Full-Length Human and Mouse cDNA Sequences. PNAS, 2002, 99(26), 16899-16903"

ORIGIN

Query Match 74.2%; Score 565.4; DB 7; Length 599;
Best Local Similarity 96.5%; Pred. No. 2.2e-158;
Matches 578; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ATGTCCGGCGGCGGAGACTTCCGGAATCCGCTGAGGAATTTCAAGCTGGTCTCTGGGG 60
Db 1 ATGTCCAGGGCGGAGACTTCCGGAATCCGCTGAGGAATTTCAAGCTGGTCTCTGGGG 60

QY 61 GAGCAAGCGTGTCAAAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTCACAA 120
Db 61 GAGCAAGCGTGTGGAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTCACAA 120

QY 121 ACCTATCAGGCAATAATTGGCATTTGATTTTATCAAAACTATGTACTTGGAGATGGA 180
Db 121 ACCTATCAGGCAATAATTGGCATTTGATTTTATCAAAACTATGTACTTGGAGATGGA 180

QY 181 ACAATCGGCTTCCGCTGTGGGATACGGCGGTGAGNACGCTCTCCGTAGCTCATTTCCC 240
Db 181 ACAATCAGGCTTCCGCTGTGGGATACGGCGGTGAGNACGCTCTCCGTAGCTCATTTCCC 240

QY 241 AGGTACATCCGTGATTTCTGCTGACGTGTAGTTTACGATATACAAATGTTAACTCA 300
Db 241 AGTTACATCCGTGATTTCTGCTGACGTGTAGTTTACGATATACAAATGTTAACTCA 300

QY 301 TTCCAGCAAACTACAAAGTGGATTTGATGATGATGATGATGATGATGATGATGATGAT 360
Db 301 TTCCAGCAAACTACAAAGTGGATTTGATGATGATGATGATGATGATGATGATGATGAT 360

QY 361 ATCAGCTAGTAGGAATAGAAAGATCTTCTGCAAGAGGCAAGTGTCAAGTGTGAGAG 420
Db 361 ATCATGCTAGTAGGAATAGAAAGATCTTCTGCAAGAGGCAAGTGTCAAGTGTGAGAG 420

QY 421 GGAGAGAGAAAGCCAAAGGCTCAATGTTTATGAACTTAGGGCAAAACTGGA 480
Db 421 GGAGAGAGAAAGCCAAAGGCTCAATGTTTATGAACTTAGGGCAAAAGCTGGA 480

QY 481 TACAATGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATCGAAAGCACA 540
Db 481 TACAATGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATCGAAAGCACA 540

QY 541 CAGGACGGAAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
Db 541 CAGGACGGAAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599


```
Query Match      69.3%; Score 528; DB 6; Length 744;
Best Local Similarity 85.7%; Pred. No. 4.3e-147;
Matches 615; Conservative 0; Mismatches 90; Indels 13; Gaps 2;

QY 57 GGGGAGCAGAAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTGA 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 GGGGAGCAGAGCGTTGGAAGACGTCCTTGATCACCAGATTCATGATGACAGTTTGA 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 117 CAACACCTATCAGGCAATTAATGGGATTCAGCTTTTATCAAAACCTATGATCTGAGGA 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 CAACACCTATCAGGCAATTAATGGGATTCAGCTTTTATCAAAACCTATGATCTGAGGA 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 177 TGAACAATTCGGCTTCGGCTGTGGATACGGCGGTGAGCAACGCTCCGTCAGCTCAT 236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 TAGAACCGTCGATTCGCAATTAATGGACACAGAGGTCAAGAGCGTTCAGGAGCTTGT 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 237 TCCAGGTACATCCGTTGATTCCTGTCAGCTGTAGTAGTTTACGATATCACAAATGTAA 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 TCTAGCTACATTCGAGACTCCACTGTGCGAGTGTGTTTATGATATCACAAATGTAA 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 297 CTCATTTCAGCAAACTACAAAGTGGATTCAGATGATGATGATGATGATGATGATGAT 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 CTCATTTCAGCAAACTACAAAGTGGATTCAGATGATGATGATGATGATGATGATGAT 302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 357 TATCATCAGCTAGTAGGAAATAGAACAGATCTTGCTGACAAAGAGCAAGTGTCAAGTGA 416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 CATCATCATGCTAGTAGGAAATAGAACAGATCTTGCTGATAGAGGCAAGTGTCAAGTGA 362
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 417 GGAGGGAGAGAGAAAGCCAAAGGCTGAATGTAGCTTTTATGAACTAGGCGCAAAAC 476
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 GGAGGGAGAGAGAAAGCCAAAGGCTGAATGTAGCTTTTATGAACTAGGCGCAAAAGC 422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 477 TGGATACAAATGTAAGAGCCTCTTTGACGCTGTAGCAGCAGCTTTGCGGGGAATGAAAG 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 AGGATACAAATGTAAGAGCCTTTTCGGCGTGTGCTGAGCTTTTACCTGGAATGAAAG 482
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 537 CACACAGGAGGAGAGAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 596
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 CACACAGGAGAGAGAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 542
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 597 ACAGTCAGGAGAGGGGTGTTCTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 ACCAGTCAATGAAGCGGCTGTTCTGCTAAATCCCTGCTGCTGCTGCTGCTGCTGCTG 602
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 657 GAA-----GCCCTTACTCTTTCAATGACTGCACTGGAATATTGGCTTGA 704
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 GAACTCACTGCTTTGGCCCCATCTTTCAATGACTGCACTGGAATATTGGCTTGA 662
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 705 CC-TTTTCCCTTCAATTAACGTTTTCGAATTCATCATTTGCTGCTGCTGCTGCTGCTG 761
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 CCTTTTCCCTTCAATTAACGTTTTCGAATTCATCATTTGCTGCTGCTGCTGCTGCTG 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
BE965054
LOCUS
DEFINITION BE965054 762 bp mRNA linear EST 14-DEC-2000
60165893191 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3886281 3',
mRNA sequence.
ACCESSION BE965054
VERSION BE965054.2 GI:11769028
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 762)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT On Oct 3, 2000 this sequence version replaced gi:10575759.
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
```

```

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC649 row: n column: 10
High quality sequence start: 43
High quality sequence stop: 728.
FEATURES
    Location/Qualifiers
        1..762
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:3886281"
            /tissue_type="large cell carcinoma, undifferentiated"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_69"
            /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.1 kb. Library constructed by Life
            Technologies."
ORIGIN
Query Match      68.9%; Score 525.2; DB 2; Length 762;
Best Local Similarity 87.4%; Pred. No. 3e-146;
Matches 626; Conservative 0; Mismatches 73; Indels 17; Gaps 4;

QY 22 GGGATCCGCTGAGGAAATTCAGCTGCTGTTCTGGGGGAGCAAGGTTGCAAAAGACA 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39 GTGAATCCGCTGAGGAAATTCAGCTGCTGTTACCTGGGGGAGCAAGGTTTGGAAAGACA 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 82 TCTTTGATCACCAGATTCAGGTATGACAGTTTTCACACACCTATCAGCGCAATATTGSC 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 99 TCTTTGATCACCAGATTCAGGTATGACAGTTTTCACACACCTATCAGCGCAATATTGSC 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 142 ATTGACTTTTATCAAAACCTATGTA--CTTGGAGGATGGA---ACAATCGGGCTTCGGC 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 ATTGACTTTTATCAAAACCTATGTAAGCTTTGGAGGATCGACACAGTAGCGATTGCAAT 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 197 TGTGGGATACGGCGGTCAGGAACGTCTCCGTAGCTCATTTCCAGGTACATCCGTGAT 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 TATGGGACACAGCAGGTCAGAGCGGTTTCAGGAGCTTGATTCCTAGCTACATTCGTGACT 278
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 257 CTGCTGCAAGCTGTAGTAGTTTACGATATCACAATGTTAACTCATTTCCAGCAAACTACAA 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 CCACGTGGCAGTTGTTGTTTATGATATCACAATGTTAACTCATTTCCAGCAAACTACAA 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 317 AGTGGATGATGATGTCAGAAACAGAAAGAGGAAGTGTATTCATCATCGCTAGTAGGAA 376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 AGTGGATGATGATGTCAGAAACAGAAAGAGGAAGTGTATTCATCATCGCTAGTAGGAA 398
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 377 ATAGAACAGATCTTGCTGACAAAGGCAAGTGTCTAGTTGAGGAGGAGAGGAAAGCCA 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 ATAAACACAGATCTTGCTGACAAAGGCAAGTGTCAATTTGAGGAGGAGAGGAAAGCCA 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 437 AAGGCTCAATGTTACGTTTATGAAACTAGGCAAAACTGGATACAAATGTAAAGCAGC 496
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 AAGAGCTGAATGTTATGTTTATGAACTAGTGCAAAAGCTGGATACAAATGTAAAGCAGC 518
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 497 TCTTTCCAGCTGTAGCAGCAGCTTTTGGCGGGAATGGAAAGCACACAGACCGGAAGCAGAG 556
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 519 TCTTTCCAGCTGTAGCAGCAGCTTTTGGCGGGAATGGAAAGCACACAGACCGGAAGCAGAG 578
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 557 AAGCATGAGTGACATAAACTGGAAGAGCTCAGGAGCAAACTAGTCAGCGAAGGGGTT 616
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 579 AAGATATGATGACATAAACTGGAAGAGCTCAGGAGCAAACTAGTCAGTGAAGAGGCT 638
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 617 GTTCTGCTACTCTCCCATGTCTCTTCAACC-----CTTCTCTCAGAACGCCCCC 665
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 639 GTTCTGCTAACTCTCCCATGTCTCTTCAACCCTTCTTCAGAAAGCTCACTGCTTTGGCCCC 698
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 666 TTACTCTTTTCATGACTGAGTGTGAATATTGGGCTTGAACCTTTTCCCTTTCATTAA 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```
Db 699 TTAATCTTTTCATTGACTGC-GTGTGACTATTGGTTGAACCTTTTCCTTCATAATAA 753
|||||
BF431157 632 bp mRNA linear EST 29-NOV-2000
7008c05.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:3573489 3'
similar to SW:RAB6_HUMAN P20340 RAS-RELATED PROTEIN RAB-6. [1] ;
mRNA sequence.
BF431157
BF431157.1 GI:11443271
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
REFERENCE
1 (bases 1 to 632)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence scop: 494.
FEATURES
Location/Qualifiers
1..632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3573489"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/notes="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
ORIGIN
Query Match 68.6%; Score 522.6; DB 2; Length 632;
Best Local Similarity 91.1%; Pred No. 1.7e-145;
Matches 555; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 1 ATGTCCGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGAGTGGTTCCTCGGG 60
Db 24 ATGTCCAGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGAGTGGTTCCTCGGG 83
Qy 61 GAGCAACGGTTGCAACACATCTTTGATCACCAGATTTCAGGTATGACAGTTTTCAGAAC 120
Db 84 GAGCAACGGTTGGAAGACATCTTTGATCACCAGATTTCAGGTATGACAGTTTTCAGAAC 143
Qy 121 ACCTATCAGGCAATAATTCGACATTTTATCAAAATCATCTACTTCGGAGATCGA 180
Db 144 ACCTATCAGGCAACAAATTCGACATTTTATCAAAATCATCTACTTCGGAGATCGA 203
Qy 181 ACAATCGGGCTTCGGCTGTGGATACGCGGGTCAGGAACTCTCCGAGCTCATTCCT 240
Db 204 ACAGTACGATTTGCAATTATGGGACACAGCAGGTCAAGAGCGGTTTCAGGAGCTTCT 263
```

```
Qy 241 AGGTACATCCGTGATTCTGCTGCAGCTGTAGTGTACGATATCACAAATGTTAACTCA 300
Db 264 AGCTACATTCGTGACTCCACTGTGGCAGTGTGTTTATGATATACAAATGTTAACTCA 323
Qy 301 TTCCAGCAAACTACAAAGTGGATTGATGTGAGAACAGAAAGAGGAAGTGATGTTATC 360
Db 324 TTCCAGCAAACTACAAAGTGGATTGATGTGAGAACAGAAAGAGGAAGTGATGTTATC 383
Qy 361 ATCAGCGTAGTAGAAATAGAACAGATCTTGCTGCAAGAGGCAAGTGTCTAGTTGAGGAG 420
Db 384 ATCATGCTAGTAGGAAATAAAAACAGATCTTCTGCAAGAGGCAAGTGTCTCAATTGAGGAG 443
Qy 421 GGAGAGAGGAAGCCAAAGGCTCAATGTTTACGTTTATTTGAACTAGGCGCAAAACTCGA 480
Db 444 GGAGAGAGGAAGCCAAAGAGCTGAATGTTTATGTTTATTTGAACTAGTGTCAAAAGCTGA 503
Qy 481 TACAATGTAAAGCAGCTCTTTTCGACGCTGTAGCAGCAGCTTTGCGGGGAATCGAAAGCACA 540
Db 504 TACAATGTAAAGCAGCTCTCTCGACGCTGTAGCAGCAGCTTTGCGGGGAATCGAAAGCACA 563
Qy 541 CAGGACGAGAGCAGAGAAAGACATGAGTGACATATAAACTGGAAAAAGCTTCAGGAGCAACA 600
Db 564 CAGGACGAGAGCAGAGAAAGATATGATTGACATATAAACTGGATAAGCCTCATGAGCAACA 623
Qy 601 GTCAGCGCAA 609
Db 624 GTCAGTGAA 632
RESULT 15
CA316338 792 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FW0-cbp-i-24-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
DEFINITION IMAGE:6812737 5', mRNA sequence.
ACCESSION CA316338
VERSION CA316338.1 GI:24534462
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
1 (bases 1 to 792)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
CDNA Library preparation: Dr. Jim Lin, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP).
Seq primer: pYX-5.
Location/Qualifiers
1..792
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6812737"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FW0"
/notes="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
```


gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACGGACACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match		67.2%;	Score 512;	DB 6;	Length 792;
Best Local Similarity		88.1%;	Pred. No. 2.8e-142;		
Matches 557;		Conservative 0;	Mismatches 75;	Indels 0;	Gaps 0;
Qy	1	ATGTCGGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTC	AAGCTGGTGTCTCTGGGG	60	
Db	65	ATGTCGGCGGCGGAGACTTCGGGAATCCGCTGAGGAAATTC	AAGCTGGTGTCTCTGGGA	124	
Qy	61	GAGCAAGCGTTGCCAAGACATCTTTGATCACCAGATTCAGGTATCAGAGTTTTCACACAC	120		
Db	125	GAGCAGAGCGTTGGAAAGACGCTTGTATCACCAGATTCAGGTATCAGAGTTTTCACACAC	184		
Qy	121	ACCTATCAGGCAATTAATTCGGCATTCGACTTTTATCAAAAATCTATGTACTTGGAGGATGGA	180		
Db	185	ACCTATCAGGCAATTAATTCGGCATTCGACTTTTATCAAAAATCTATGTACTTGGAGGATGGA	244		
Qy	181	ACAATCGGCGTTCGGCTGTGGATACGGCGGTGAGGAACGTCCTCGTAGCCTCATTTCCC	240		
Db	245	ACCGTCCGATTGCAATTTATGGGACACAGCAGGCTCAAGCGGTTTCAGGAGCTTGATTCT	304		
Qy	241	AGGTACATCCGATTTCTGCTCAGCTGTAGTGTACGATATCACAATGTAACTCA	300		
Db	305	ACCTACATTCGAGACTCCACTGTGGCAGTTGTTTATGATATCACAATGTAACTCA	364		
Qy	301	TTCCAGCAAACTACAAAGTGGATGTATGATGTACAGAACAGAAAGAGGAGTGTATTC	360		
Db	365	TTCCAGCAAACTACAAAGTGGATGTATGATGTACAGAACAGAGAGGAGTGTATTCATC	424		
Qy	361	ATCAGCTAGTAGGAAATAGAACAGATCTTGTCTGACAGAGGCAAGTGTCACTTGGAGAG	420		
Db	425	ATCATGCTAGTAGGAAATAGAACAGATCTTGTCTGATAGAGGCAAGTGTCACTTGGAGAG	484		
Qy	421	GGAGAGGAAAGCCAAAGGCTGAATGTTACGTTTATTGAACTAGGCAAAACTGGA	480		
Db	485	GGAGAGGAAAGCCAAAGGCTGAATGTTATGTTTATTGAAACCCAGTCAAAAGCAGGA	544		
Qy	481	TACAATGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTTCCGGGAATGGAAGCACA	540		
Db	545	TACAATGTAAAGCAGCTTTTCCGGCGTGTGTCTGCAGCTTTTACCTGGAATGGAAGCACA	604		
Qy	541	CAGGACGGAAGCAGAGAACATAGTGTACATAAACTGGAAGAGCCTCAGGAGCAACA	600		
Db	605	CAGGACGGAAGCAGAGAACATAGTGTACATAAACTGGAAGAGCCTCAGGAGCAACA	664		
Qy	601	GTCAGGAGAGGGGTGTTCTCTGCTACTCTCC	632		
Db	665	GTCATGAAGCGCGCTGTTCTGCTAATCCCC	696		

Search completed: April 25, 2005, 11:40:30
Job time : 3105 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2005, 22:22:10 ; Search time 3301 Seconds
(without alignments)

3728.453 Million cell updates/sec

Title: US-09-441-857-12

Perfect score: 1304

Sequence: 1 MSAGGDFGNLKRKFLVFLG.....NLFPSTLTFNCSSLLPVSWR 254

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09441857/runat_22042005_155508_24747/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09441857 @CGN 1 1 5600 @runat_22042005_155508_24747 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.ev.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1304	100.0	765	9	CR533469 Homo sapi
2	1304	100.0	840	9	AF309646 Homo sapi
3	1304	100.0	3084	6	AX086357 Sequence
4	1299	99.6	198961	9	AC079776 Homo sapi

C	5	1299	99.6	349981	6	AX647880	Sequence
C	6	1299	99.6	349989	6	AX647177	Sequence
C	7	1293	99.2	165411	9	AC018865	Homo sapi
C	8	1288	98.8	195280	9	AC073869	Homo sapi
C	9	1261	96.7	242865	2	AC120839	Pan trogl
	10	1255	96.2	184706	9	AC120781	Homo sapi
	11	1147	88.0	3070	9	BC003617	Homo sapi
	12	1147	88.0	3100	9	BC068486	Homo sapi
	13	1147	88.0	3120	9	AK057157	Homo sapi
	14	1138	87.3	3745	6	AB447780	Sequence
	15	1137	87.2	1770	10	BC019118	Mus muscu
	16	1089	83.5	1950	10	AB041575	Mus muscu
	17	1062.5	81.5	1170	5	CR390639	Gallus ga
	18	1012.5	77.6	844	5	AF540409	Melagris
	19	1000	76.7	2527	10	BC026915	Mus muscu
	20	982	75.3	739	9	AF119836	Homo sapi
	21	960	73.6	683	9	AF130122	Homo sapi
	22	959	73.5	740	6	AX771444	Sequence
	23	959	73.5	740	9	HUMRAB6A	M28212 Homo sapien
	24	959	73.5	2099	5	BC046683	Xenopus l
	25	953	73.1	627	9	AF198616	Homo sapi
	26	953	73.1	627	9	AF498941	Homo sapi
	27	948	72.7	627	9	CR541984	Homo sapi
	28	946.5	72.6	1461	5	BC064230	Xenopus t
	29	944	72.4	627	9	AF130986	Homo sapi
	30	944	72.4	627	9	AF498939	Homo sapi
	31	897.5	68.8	4768	10	BC060618	Mus muscu
	32	892	68.4	627	9	AF498940	Homo sapi
	33	892	68.4	627	9	BT007263	Homo sapi
	34	892	68.4	1261	9	BC002510	Homo sapi
	35	892	68.4	1266	9	AF166492	Homo sapi
	36	888	68.1	5395	9	BC078662	Homo sapi
	37	884	67.8	1589	6	BD131169	Human tra
	38	883	67.7	573	9	AF124200	Homo sapi
	39	878	67.3	1652	5	BC074238	Xenopus l
	40	869	66.6	2005	5	BC080215	Danio rer
	41	864.5	66.3	172962	10	AL713865	AL713865 Mouse DNA
	42	860	66.0	2498	5	BC058059	Danio rer
	43	846.5	64.9	1157	3	D84314	Drosophila
	44	846.5	64.9	1188	6	C0586572	Sequence
	45	846.5	64.9	1857	3	AY060261	Drosophil

ALIGNMENTS

RESULT 1	CR533469	765 bp	mRNA	linear	PRI 22-JUN-2004
LOCUS	CR533469	Homo sapiens full open reading frame cDNA clone RZPD0834C0517D for gene RAB6C, RAB6C, member RAS oncogene family; complete cds, incl. stopcodon.			
DEFINITION	CR533469.1	GI:49065363	Full ORF shuttle clone, Gateway(TM), complete cds.		
ACCESSION	CR533469	Homo sapiens (human)			
VERSION	CR533469.1	Homo sapiens			
KEYWORDS	Full ORF shuttle clone, Gateway(TM), complete cds.	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
SOURCE	1 (bases 1 to 765)	Ebert, L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.			
ORGANISM	Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201)	Unpublished			
REFERENCE	2 (bases 1 to 765)	Ebert, L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.			
AUTHORS	Direct Submission	Submitted (21-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany			
JOURNAL	Submitted (21-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany	RZPD; RZPD0834C0517D, ORFNO 2784 www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPD0834C0517D RZPDLIB; Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No. 834			
COMMENT					

www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834
 www.rzpd.de/products/orfclones/
 Contact: Inge Ariart
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 100
 Fax: +49 30 32639 111
 www.rzpd.de

This clone is available from RZPD;
 contact RZPD (customer.service@rzpd.de) for further information.
 This CDS clone is a part of a collection of human full ORF clones
 generated by RZPD.
 This CDS has been cloned incl. stopcodon.
 The CDS has been inserted into pDONR201 via a BP Clonase(TM)
 reaction. Additional sequence has been added in front of the start
 codon: att..AAAAA GCA GGC (ATG).
 After the stop codon 3' UTR sequence is present in front of the 3'
 att site (ACCCAGCTTTCTT).
 Compared to the reference sequence NM_032144 (gi14149798) we did
 not find any amino acid exchanges.
 Clone distribution: http://www.rzpd.de/products/orfclones/.

FEATURES

source

```

1..765
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPD0834C0517D"
/clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
/lab_host="DH10B"
/notes="Vector: pDONR201, Site_1: attP1; Site_2: attP2"
gene
1..765
/feature="RAB6C"
CDS
1..765
/feature="RAB6C"
/codon_start=1
/protein_id="CAG38500.1"
/db_xref="GI:49065364"
/translation="MSAGDFGNPLKFKLVFLGEQSVAKTSLITFRYDSFDNTYQA
IIGIDFSLKMYLEDGTIGRLMDTAGQRLRSILPIRYIRDSAAAAYVVDITNVNSFQ
QTTWIDVTRGSDVITLVGNRTDLADKQVSEVEGERKAKGLNVTFIETRAKTG
YNVKQLFRVAAALPGMESTQDGSREDMSDIKLEKPEQTVSEGGSCSYSPMSSSTLP
QKPPYSFIDCSVNIQLNLPSSLTPVSWR"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 4,1e-118 Length: 765
Score: 1304.00 Matches: 254
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

US-09-441-857-12 (1-254) x CR533469 (1-765)

```

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 1 ATGTCGCGGGGAGACTTCGGGAATCCGCTAGGAAATTCAGCTGGTCTCTCGGG 60

Qy 21 GluInSerValAlaLysThrSerLeuLeThrArgPheArgTyrAspSerPheAspAsn 40
Db 61 GAGCAACGCTGTCAGACACATCTTGATCACCAGATTCAGGTATGACAGTTTTGACAAC 120

Qy 41 ThrTyrGlnAlaLysLeuGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 121 ACCTATCAGGCAATTAATGGCAATTCACATTTTATCAAAACATATGATCTTGAGAGTGA 180

Qy 61 ThrileGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuilePro 80
Db 181 ACAATCGGCTTCGGCTGTGGGATACCGGGGTGAGAACCTCTCCGTAGCCTCATTTCCC 240

Qy 81 ArgTyrileArgAspSerAlaAlaValValValTyrAspilleThrAsnValAsnSer 100
Db 241 AGGTACATCCGTGATTCTGCTGCAGCTGTAGTAGTTTACGATATACAAATGTTAACTCA 300

```

```

Qy 101 PheGlnGlnThrThrLysTyrPilleAspAspValArgThrGluArgGlySerAspValille 120
Db 301 TTCCAGCAAACTACAAAGTGGATTGATGATCTCAGAACAGAAAGAGGAAGTGTATGTTATC 360

Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 361 ATCAGCTAGTAGGAATATAGAACAGATCTTGCTGCAAGAGGCAAGTGTCAAGTTGAGAG 420

Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheilleGluThrArgAlaLysThrGly 160
Db 421 GGAGAGAGGAAGCAAAAGGGCTGAATGTTACGTTTATGAACTAGGCAAAACTGGA 480

Qy 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLalLeuProGlyMetGluSerThr 180
Db 481 TACAATGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCGGGAAATGGAAGACA 540

Qy 181 GlnAspGlySerArgGluAspMetSerAspLysLeuGluLysProGlnGluGlnThr 200
Db 541 CAGGACGAGACAGAGAGACATGAGTACATAAACTGGAAGACCTTCAGGAGCAACA 600

Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 601 GTCAGCGAAGGGGTTGTTCTGCTACTCTCCCATGTCATCTTCAACCTTCTCAGAG 660

Qy 221 ProProTyrSerPheilleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeu 240
Db 661 CCCCCTTACTCTTTTCAATGACTGCAGTGTGAATATTCGCTTGAACCTTTTCCCTTCATTA 720

Qy 241 IleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 721 ATAACGTTTTGCAATTCATTCGCTGCTCTCGTGGAGG 762

```

RESULT 2

AF309646

LOCUS

AF309646

DEFINITION

Homo sapiens Rab6-like protein WTH3 mRNA, complete cds.

ACCESSION

AF309646

VERSION

AF309646.1

KEYWORDS

GI:23820935

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Shan,J., Yuan,L., Budman,D.R. and Xu,H.P.

WTH3, a new member of the Rab6 gene family, and multidrug

resistance

Biochim. Biophys. Acta 1589 (2), 112-123 (2002)

22003393

PUBMED

12007787

REFERENCE

2 (bases 1 to 840)

Shan,J.D., Yuan,L.M. and Xu,H.P.

Direct Submission

TITLE

Submitted (29-SEP-2000) Molecular Oncology, North Shore University

Hosptial, 350 Community Drive, Manhasset, NY 11030, USA

JOURNAL

Location/Qualifiers

FEATURES

source

1..840

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2q31"

48..812

/notes="similar to GTP-binding protein"

/codon_start=1

/product="Rab6-like protein WTH3"

/protein_id="AA39685.1"

/db_xref="GI:23820936"

/translation="MSAGDFGNPLKFKLVFLGEQSVAKTSLITFRYDSFDNTYQA

IIGIDFSLKMYLEDGTIGRLMDTAGQRLRSILPIRYIRDSAAAAYVVDITNVNSFQ

QTTWIDVTRGSDVITLVGNRTDLADKQVSEVEGERKAKGLNVTFIETRAKTG

YNVKQLFRVAAALPGMESTQDGSREDMSDIKLEKPEQTVSEGGSCSYSPMSSSTLP


```

Qy      241 1leThrPheCysAsnSerLeuLeuProValSerTtpAtg 254
Db      1176 ATACGTTTTCATTCATTCATTCGCTGCTGCTGAGG 1217

RESULT 4
AC079776      198961 bp      DNA      linear      PRI 22-FEB-2002
LOCUS      Homo sapiens BAC clone RP11-315H12 from 2, complete sequence.
DEFINITION      AC079776
ACCESSION      AC079776
VERSION      AC079776.5 GI:18482321
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 198961)
AUTHORS      Sulston, J.E. and Waterston, R.
TITLE      Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE      99063792
PUBMED      9847074
REFERENCE      2 (bases 1 to 198961)
AUTHORS      Shah, N. and Kozlowicz, A.
TITLE      The sequence of Homo sapiens BAC clone RP11-315H12
JOURNAL      Unpublished (2001)
REFERENCE      3 (bases 1 to 198961)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (10-SEP-2000) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
REFERENCE      4 (bases 1 to 198961)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (03-FEB-2002) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
REFERENCE      5 (bases 1 to 198961)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (06-FEB-2002) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
REFERENCE      6 (bases 1 to 198961)
AUTHORS      Waterston, R.
TITLE      Direct Submission
JOURNAL      Submitted (22-FEB-2002) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              On Feb 3, 2002 this sequence version replaced gi:18042404.
              ----- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu/gsc
              Contact: sapiens@watson.wustl.edu
              ----- Summary Statistics
              -----
              Center project name: H_NH0315H12
              -----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.

```

McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-228P15, 2000 bp overlap; the clone sequenced to the right is RP11-32C20. Actual start of this clone is at base position 129432 of RP11-228P15; actual end is at base position 198961 of RP11-315H12.

Data from AC079450 was used to finish AC079776.

FEATURES	Location/Qualifiers
source	1..198961
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-315H12"
	/clone_lib="RPCI-11"
repeat_region	216..817
	/rpt_family="L1"
repeat_region	818..1399
	/rpt_family="MaLR"
repeat_region	1403..2150
	/rpt_family="MaLR"
repeat_region	2148..2492
	/rpt_family="L1"
repeat_region	2534..3082
	/rpt_family="L1"
repeat_region	3106..3742
	/rpt_family="MaLR"
repeat_region	3763..4054
	/rpt_family="Alu"
repeat_region	4084..4452
	/rpt_family="MaLR"
repeat_region	4456..4666
	/rpt_family="Alu"
repeat_region	4671..4753
	/rpt_family="L1"
repeat_region	4739..5052
	/rpt_family="L1"
repeat_region	5057..5148
	/rpt_family="L1"
repeat_region	5149..5262
	/rpt_family="Alu"
repeat_region	6702..6877
	/rpt_family="Alu"
repeat_region	6878..7812
	/rpt_family="L1"
repeat_region	7812..8134
	/rpt_family="L1"
repeat_region	8135..8434
	/rpt_family="Alu"
repeat_region	8435..9553
	/rpt_family="L1"
repeat_region	9554..9574
	/rpt_family="AT-rich"
repeat_region	9579..9658
	/rpt_family="L1"
repeat_region	9659..9961
	/rpt_family="Alu"
misc_feature	9936..9958

/note="old seq 1369: original length 744.802 splitted into
3 new seqs-new seq 1369: from 000.001 to 349.980-new seq
2072: from 300.001 to 649.980-new seq 2073: from 600.001
to 744.802"

ORIGIN

Alignment Scores:
Pred. No.: 4.11e-114 Length: 349981
Score: 1299.00 Matches: 253
Percent Similarity: 99.61% Conservative: 0
Best Local Similarity: 99.61% Mismatches: 1
Query Match: 99.62% Indels: 0
DB: 6 Gaps: 0

US-09-441-857-12 (1-254) x AX647780 (1-349981)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 45699 ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTTCCTGGGG 45640

Qy 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAspAsn 40
Db 45639 GAGCAAGCGTTGCCAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGGACAAC 45580

Qy 41 ThrTyrGlnAlaIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 45579 ACCTATCAGGCATTAATTTGGCATTGACCTTTTATCAAAACATATGATCTTGAGGATGGA 45520

Qy 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 45519 ACAATCGGCTTCGCGTGGGATACGGCGGGTCAGGAACGCTCCGCTAGCTCATTTCCC 45460

Qy 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
Db 45459 AGGTACATCCGCTGATTCTGCTGCACCTGTAGTAGTTTACGATATCACAAATGTTAACTCA 45400

Qy 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 45399 TTCCAGCAAACTACAAATGGATGATGATGTCAGAACAGAAAGGAAAGTGTATTC 45340

Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 45339 ATCAGCTAGTAGGAAATAGAACATCTTCTGACAGAGGCAAGTGTGAGTGGAG 45280

Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 45279 GGAGAGAGAAAGCCAAAGGGCTGAATGTTACGTTTATTGAAACTAGGCGCAAAAGCTGGA 45220

Qy 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 45219 TACAATGTAAAGCAGCTCTTTGCGAGTGTAGCAGCAGCTTTGCCGGGAATGGAAGCACA 45160

Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnGlnThr 200
Db 45159 CAGGACGGAACAGAGAGACATGAGTCACATAAACTGGAAGCCCTCAGGACCAACA 45100

Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerThrLeuProGlnLys 220
Db 45099 GTCAGCAAGGGGGTGTGTTCTCTCTACTCTCCATGTCTTCAACCCCTTCTCAGAAG 45040

Qy 221 ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeu 240
Db 45039 CCCCCCTTACTCTTCAATGACGTGAGTGTGAATATGCTGCTTGAACCTTTTCCCTTCATTA 44980

Qy 241 IleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 44979 ATACGTTTTCGAATTCATCATTGCTGCTGTCTCTCTGCGAGG 44938

RESULT 6
AX647177/c
LOCUS
DEFINITION Sequence 1369 from Patent EP1270724.
ACCESSION AX647177

VERSION
KEYWORDS
SOURCE
ORGANISM

AX647177.1 GI:28800410

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.

Guanosine triphosphate-binding protein coupled receptors

Patent: EP 1270724-A 1369 02-JAN-2003;

National Institute of Advanced Industrial Science and Technology

(JP) ; Center for Advanced Science and Technology Incubation, Ltd.

FEATURES

Location/Qualifiers

source

1. .349989

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="old seq 1369: original length 744.802 splitted into

3 new seqs-new seq 1369: from 000.001 to 349.980 *-new seq

2072: from 300.010 to 649.990 *-new seq 2073: from 600.011

to 744.802 *"

join(201. .246,25640. .25677,27078. .27094,141192. .141769,

159571. .159606,174525. .174575,234891. .235013,

235514. .235560,279677. .279729)

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD70100.1"

/db_xref="GI:28800411"

/translation="MAASISEASDAMDPTGFDVAELPAAVATILGALLAAATLAVS

PLLVITLRNQRQRPHVLLPANILLDLAYILLHMLISSSSIGSGWELGRMACGILT

DAYPACTSTILSFTALVTHYLAVHPLRYLSFMSHGAAKVAVALWLVAACRPTPL

ILWSKQDAQEQQASVILPPSGTQPGCLLVITVTYSILCVLICTALIANCFWR

IYAEATSVAGSTQRLRHERDPSGRRSKAPLPFGQLSSSGSTFRPQLRVGLAHL

PIQDEDVSSRLDGLVLCRLPLTVMLPVPFELSPQAAAPCAMFWLQRCV"

ORIGIN

Alignment Scores:

Pred. No.: 4.11e-114 Length: 349989

Score: 1299.00 Matches: 253

Percent Similarity: 99.61% Conservative: 0

Best Local Similarity: 99.61% Mismatches: 1

Query Match: 99.62% Indels: 0

DB: 6 Gaps: 0

US-09-441-857-12 (1-254) x AX647177 (1-349989)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20

Db 345708 ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTTCCTGGGG 345649

Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40

Db 345648 GAGCAAGCGTTGCCAAGACATCTTTCATCACCAGATTCAGGTATGACAGTTTGGACAAC 345589

Qy 41 ThrTyrGlnAlaIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60

Db 345588 ACCTATCAGCAATAATTCGCATTGACCTTTTATCAAAAAATATGTACTTGGAGGATGGA 345529

Qy 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80

Db 345528 ACAATCGGCTTCGCTGTGGATACGCGGGTCAGGAACGTCCTCCGTAGCTCATTTCCC 345469

Qy 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100

Db 345468 AGGTACATCCGCTGATTCTGCTGAGCTGTAGTAGTTTACGATATCACAATGTTAACTCA 345409

Qy 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120

Db 345408 TTCAGCAAACTCAAAAGTGGATTGATGTGATGTGACAAACAGAAAGAGTGTGTTATC 345349

Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140


```

Db      345348 ATCACCTAGTAGGAATAAGAACAGATCTTGTGACAGGCAAGTGTGAGTGTGAGGAG 345289
QY      141 GYGluArgLysGlyLeuAsnValThrPheIlecluThrArgAlaLysThrGly 160
Db      345288 GGNAGAGGAGNAGCCAAAGGGCTGAATGTACGTTTATTGAACTAGGCAAAAGCTGGA 345229
QY      161 TyrAsnValGlyGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
Db      345228 TACAATGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAGCACA 345169
QY      181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db      345168 CAGGACGGAAGCAGAGAAACATGAGTGACATAAAACTGGAAGAGCTCAGGAGCAACA 345109
QY      201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerThrLeuProGlnLys 220
Db      345108 GTCAGGAGAGGGGGTGTTCCTGCTACTCTCCATGTCATCTTCAACCCCTTCTCAGAAG 345049
QY      221 ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeu 240
Db      345048 CCCCCTTACTCTTTTCACTGACGTCAGTGTGAATATTGGCTTGAACCTTTTCCCTTCATTA 344989
QY      241 IleThrPheCysAsnSerSerLeuLeuProValSerTyrArg 254
Db      344988 ATAACGTTTTTGCAATTCAATTCATTCGTCGCTGCTGTCGTGAGG 344947

RESULT 7
AC018865      165411 bp      DNA      linear      PRI 23-NOV-2002
LOCUS      Homo sapiens BAC clone RP11-32C20 from 2, complete sequence.
AC018865
ACCESSION      AC018865
VERSION      AC018865.4 GI:25189068
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 165411)
AUTHORS      Sulston, J.E. and Waterston, R.
TITLE      Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE      99063792
PUBMED      9847074
REFERENCE      2 (bases 1 to 165411)
AUTHORS      Kim, K., Haglund, K., Tomlinson, C., Mangiapanello, L. and Dignan, G.
TITLE      The sequence of Homo sapiens BAC clone RP11-32C20
JOURNAL      Unpublished (2001)
REFERENCE      3 (bases 1 to 165411)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (21-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE      4 (bases 1 to 165411)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (13-SEP-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE      5 (bases 1 to 165411)
AUTHORS      Waterston, R.
TITLE      Direct Submission
JOURNAL      Submitted (23-NOV-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 23, 2002 this sequence version replaced gi:22830544.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0032C20
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

There is an unresolved tandem repeat from base 115267 to 124390.

Polymorphisms have been identified between AC018804 and this sequence.

Data from AC018804 was used to finish this clone.

This sequence is the entire insert of the clone.

FEATURES	source
Location/Qualifiers	1..165411
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-32C20"
	/clone_lib="RPCI-11"
	487..1684
repeat_region	/rpt_family="L1"
	1678..1885
repeat_region	/rpt_family="L1"
	1890..1986
repeat_region	/rpt_family="L1"
	1987..2289
repeat_region	/rpt_family="Alu"
	2294..2380
repeat_region	/rpt_family="GA-rich"
	2392..2699
repeat_region	/rpt_family="L1"
	2905..3262
repeat_region	/rpt_family="Alu"
	3284..3395
repeat_region	/rpt_family="T-rich"
	3601..3674
repeat_region	/rpt_family="MER105"
	3726..3845
repeat_region	/rpt_family="Alu"
	3846..3869
repeat_region	/rpt_family="(TAA)n"
	3870..4042
repeat_region	/rpt_family="Alu"
	4064..4364

repeat_region	14431..14726	/rpt_family="Alu"	14726	14431	repeat_region	14431..14726	/rpt_family="Alu"	14726	14431
repeat_region	14754..14809	/rpt_family="MIR"	14809	14754	repeat_region	14754..14809	/rpt_family="MIR"	14809	14754
repeat_region	15272..15400	/rpt_family="CT-rich"	15400	15272	repeat_region	15272..15400	/rpt_family="CT-rich"	15400	15272
repeat_region	15425..15458	/rpt_family="(TTCTC)n"	15458	15425	repeat_region	15425..15458	/rpt_family="(TTCTC)n"	15458	15425
repeat_region	15465..15650		15650	15465	repeat_region	15465..15650		15650	15465
repeat_region	15678..15704	/rpt_family="AT-rich"	15704	15678	repeat_region	15678..15704	/rpt_family="AT-rich"	15704	15678
repeat_region	15781..15781	/rpt_family="AT-rich"	15781	15781	repeat_region	15781..15781	/rpt_family="AT-rich"	15781	15781
repeat_region	15822..16005	/rpt_family="Alu"	16005	15822	repeat_region	15822..16005	/rpt_family="Alu"	16005	15822
repeat_region	16067..16146	/rpt_family="MER1_type"	16146	16067	repeat_region	16067..16146	/rpt_family="MER1_type"	16146	16067
repeat_region	16196..16487	/rpt_family="Alu"	16487	16196	repeat_region	16196..16487	/rpt_family="Alu"	16487	16196
repeat_region	16515..16634	/rpt_family="MER1_type"	16634	16515	repeat_region	16515..16634	/rpt_family="MER1_type"	16634	16515
repeat_region	16755..17044	/rpt_family="Alu"	17044	16755	repeat_region	16755..17044	/rpt_family="Alu"	17044	16755
repeat_region	17349..17641	/rpt_family="Alu"	17641	17349	repeat_region	17349..17641	/rpt_family="Alu"	17641	17349
repeat_region	17644..17932	/rpt_family="Alu"	17932	17644	repeat_region	17644..17932	/rpt_family="Alu"	17932	17644
repeat_region	17948..18108	/rpt_family="BC200"	18108	17948	repeat_region	17948..18108	/rpt_family="BC200"	18108	17948
repeat_region	18294..18589	/rpt_family="Alu"	18589	18294	repeat_region	18294..18589	/rpt_family="Alu"	18589	18294
repeat_region	18824..19036	/rpt_family="L1"	19036	18824	repeat_region	18824..19036	/rpt_family="L1"	19036	18824
repeat_region	19037..19342	/rpt_family="Alu"	19342	19037	repeat_region	19037..19342	/rpt_family="Alu"	19342	19037
repeat_region	19343..19404	/rpt_family="L1"	19404	19343	repeat_region	19343..19404	/rpt_family="L1"	19404	19343
repeat_region	19405..19444	/rpt_family="Alu"	19444	19405	repeat_region	19405..19444	/rpt_family="Alu"	19444	19405
repeat_region	19445..19472	/rpt_family="(TTTTG)n"	19472	19445	repeat_region	19445..19472	/rpt_family="(TTTTG)n"	19472	19445
repeat_region	19473..19748	/rpt_family="Alu"	19748	19473	repeat_region	19473..19748	/rpt_family="Alu"	19748	19473
repeat_region	19749..19903	/rpt_family="L1"	19903	19749	repeat_region	19749..19903	/rpt_family="L1"	19903	19749
repeat_region	10276..10381	/rpt_family="MIR"	10381	10276	repeat_region	10276..10381	/rpt_family="MIR"	10381	10276
repeat_region	10488..10516	/rpt_family="(A)n"	10516	10488	repeat_region	10488..10516	/rpt_family="(A)n"	10516	10488
repeat_region	10517..10806	/rpt_family="Alu"	10806	10517	repeat_region	10517..10806	/rpt_family="Alu"	10806	10517
repeat_region	10918..10942	/rpt_family="(GAA)n"	10942	10918	repeat_region	10918..10942	/rpt_family="(GAA)n"	10942	10918
repeat_region	11350..11603	/rpt_family="Alu"	11603	11350	repeat_region	11350..11603	/rpt_family="Alu"	11603	11350
repeat_region	11614..11738	/rpt_family="MIR"	11738	11614	repeat_region	11614..11738	/rpt_family="MIR"	11738	11614
repeat_region	11961..12196	/rpt_family="MIR"	12196	11961	repeat_region	11961..12196	/rpt_family="MIR"	12196	11961
repeat_region	12222..12355	/rpt_family="CR1"	12355	12222	repeat_region	12222..12355	/rpt_family="CR1"	12355	12222
repeat_region	12365..12448	/rpt_family="MIR"	12448	12365	repeat_region	12365..12448	/rpt_family="MIR"	12448	12365
repeat_region	12449..12747	/rpt_family="Alu"	12747	12449	repeat_region	12449..12747	/rpt_family="Alu"	12747	12449
repeat_region	12802..12853	/rpt_family="L1"	12853	12802	repeat_region	12802..12853	/rpt_family="L1"	12853	12802
repeat_region	12855..13172	/rpt_family="Alu"	13172	12855	repeat_region	12855..13172	/rpt_family="Alu"	13172	12855
repeat_region	13199..13700	/rpt_family="L1"	13700	13199	repeat_region	13199..13700	/rpt_family="L1"	13700	13199
repeat_region	13701..13700		13700	13701	repeat_region	13			

AC073869/c
 LOCUS AC073869 195280 bp DNA linear PRI 09-JAN-2002
 DEFINITION Homo sapiens BAC clone RP11-109E12 from 2, complete sequence.
 ACCESSION AC073869
 VERSION AC073869.5 GI:15638756
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 195280)
 Suleton, J.E. and Waterston, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 99063792
 9847074
 2 (bases 1 to 195280)
 Bielicki, L., Maupin, R. and Meyer, R.
 The sequence of Homo sapiens BAC clone RP11-109E12
 Unpublished (2001)
 3 (bases 1 to 195280)
 Waterston, R.H.
 Direct Submission
 Submitted (30-JUN-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 195280)
 Waterston, R.H.
 Direct Submission
 Submitted (18-SEP-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 195280)
 Waterston, R.
 Direct Submission
 Submitted (09-JAN-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Sep 18, 2001 this sequence version replaced gi:14209794.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.edu
 ----- Summary Statistics

 Center project name: H_NH0109E12

 NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.
 This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.
 MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D.
 McPherson, Department of Genetics, Washington University, St. Louis
 MO. For additional information about the map position of this
 sequence, see <http://genome.wustl.edu/gsc>
 SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
 Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong

and coworkers at the Roswell Park Cancer Institute
 (<http://bacpac.med.buffalo.edu>)
 VECTOR: pBACe3.6
 NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is RP11-433A19. Actual start of
 this clone is at base position 1 of RP11-109E12; actual end is at
 base position 195280 of RP11-109E12.
 Data from AC022629 was used to finish AC073869.
 FEATURES
 Location/Qualifiers
 1..195280
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-109E12"
 /clone_lib="RPCI-11"
 1..167
 /note="match to EST AA904412 (NID:g3039535) ok07e01.s1"
 82..96
 /note="similar to Homo sapiens EST AV651649
 (NID:g9872663)"
 298..778
 /note="match to EST AW293618 (NID:g6700254)"
 408..431
 /note="similar to Homo sapiens EST AV651649
 (NID:g9872663)"
 503..670
 /note="match to EST AI352084 (NID:g4089290) qr10b01.x1"
 651..1100
 /note="match to EST AI138580 (NID:g3644552) qd85b12.x1"
 695..710
 /note="similar to Homo sapiens EST BE791748
 (NID:g10212946)"
 1698..2014
 /note="similar to Homo sapiens EST N30408 (NID:g1148928)
 YJ13C09.s1"
 2015..2410
 /note="similar to Homo sapiens EST AV651649
 (NID:g9872663)"
 2253..2732
 /note="match to EST AW340014 (NID:g6836640) hc91e04.x1"
 2338..3119
 /note="similar to Homo sapiens EST AUL139340
 (NID:g11000861)"
 2340..2885
 /note="similar to Bos taurus EST BE722624 (NID:gl0123920)"
 3154..3941
 /note="similar to Canis familiaris EST B1430421
 (NID:g15207653)"
 3236..3645
 /note="similar to Mus musculus EST AA617192 (NID:g2504397)
 VJ79a09.r1"
 4375..4376
 /note="similar to Homo sapiens EST AI677857 (NID:g4888039)
 wd34c04.x1"
 4427..4468
 /rpt_family="(TTTA)n"
 4440..4738
 /rpt_family="Alu"
 4835..4973
 /note="similar to Homo sapiens EST AI677857 (NID:g4888039)
 wd34c04.x1"
 4892..5127
 /note="similar to Homo sapiens EST BF511460
 (NID:g11594758)"
 4899..5106
 /note="match to EST AW804021 (NID:g7855891)"
 4974..5130
 /note="similar to Homo sapiens EST AI677857 (NID:g4888039)
 wd34c04.x1"
 5205..5387
 misc_feature

Qy	1	MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly	20
Db	128482	ATGTCCGGCGCGCGACCTCGGAAATCCGCTGAGAAATTCAGCTGGTCTCTGGGG	128423
Qy	21	GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn	40
Db	128422	GAGCAAAAGCGTGTCAAAAGACATCTTTGATCATCCAGATTTCATGTATGACAGTTTTTCACAAAC	128363
Qy	41	ThrTyrGlnAlaIleIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly	60
Db	128362	ACCTATCAGGCACAAATTGGCATTGACATTTTATCAAAAACCTATGTACTTGGAGGATGGA	128303
Qy	61	ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro	80
Db	128302	ACAATCGGGCTTCGGCTGTGGGATACGGCGGTACAGGAACGTCCTCGTAGCCCTCATTCOC	128243
Qy	81	ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer	100
Db	128242	AGGTACATCCGTGATTCGTCTGCAGCTGTAGTAGTTTACGATATCACAAATGTTAACTCA	128183
Qy	101	PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle	120
Db	128182	TTCCAGCAAACTACAAAGTGATGATGATGTGCAGACAGAAAGAGGAAGTGTATCTATC	128123
Qy	121	IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu	140
Db	128122	ATCACTCTAGTAGGAAATAGAACAGATCTTGCTGCAAGAGGCAAGTAGTGTCAATCTCAGGAG	128063
Qy	141	GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly	160
Db	128062	GGAGAGGAGGAAGCCAAAGGGCTGAACTGTAGGTTTATTGAAACCTAGGCGCAAAAGCTGGA	128003
Qy	161	TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr	180

```
/db_xref="taxon:9598"
/clone="rp43-28ml17"
/clone_lib="RP43 - 43 Male Chimpanzee BAC Library"

ORIGIN

Alignment Scores:
Pred. No.:      3.57e-110      Length:      184706
Score:          1255.00      Matches:      245
Percent Similarity: 96.85%      Conservative: 1
Best Local Similarity: 96.46%      Mismatches: 8
Query Match:      96.24%      Indels:      0
DB:              9          Gaps:      0

US-09-441-857-12 (1-254) x AC120781 (1-184706)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 86155 ATGTCGGCGGCGGCGACCTCGGAATCCGCTGAGGAAATTCAGAGCTGGTCTCTGGG 86214

Qy 21 GluGlnSerValAlaLysThrSerLeuileThrArgPheA-gTyrAspSerPheAspAsn 40
Db 86215 GAGCAAAACGGTTGCAAGACATCTTTGATCACCAGATTTCATGATGACAGTTTTCACAAC 86274

Qy 41 ThrTyrGlnAlaIleIleGlyLeAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 86275 ACCTATCAGGCAACAATTGGCATTGACATTTTATCAAAAACATGATGACTTGGAGATGGA 86334

Qy 61 ThrLeGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 86335 ACAATCGGCTTCGGCTCGGATACGGCGGTCAGGAACGTCCTCGTAGCTCATTCCTCC 86394

Qy 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
Db 86395 AGGTACATCCGTGATCTCTGCGAGCTGTAGTAGTTACGATATACAAAATGTTAACTCA 86454

Qy 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 86455 TTCAGCAAACTACAAAGTGATGTATGATGTCTCAGACAGAAAGAGGAGTGTATCTC 86514

Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 86515 ATCACTCTAGTAGGAATAGAACAGATCTTCTGCAAGAGGCAAGTGTCAATTGAGGAG 86574

Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 86575 GGAGAGAGAAAGCAAGAGGGCTGAATGTAGTTTATTGAAACTAGGGCAAAAAGCTGGA 86634

Qy 161 TyrAsnValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr 180
Db 86635 TACATGTAAAGCAGCTCTTCGACGTGTAGCAGCAGCTTCGCGGGNATGGAAAGCACA 86694

Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThr 200
Db 86695 CAGGACAGAACGACAGAGACATGATGACATAAACTGGAAAGACCTCAGGACCAACCA 86754

Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGluLys 220
Db 86755 GTCACGCAAGGGGGTGTCTCTGCTACTCTCCCATGTCATCTTCAACCCCTTCCTCAGAAG 86814

Qy 221 ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeu 240
Db 86815 CCCCCTACTCTTTCATGATGATGAGTGTGAATATGGCTTGAACCTTTCCCTTCATTA 86874

Qy 241 IleThrPheCysAsnSerSerLeuProValSerTrpArg 254
Db 86875 ATAACGTTTTGCAATTTCATCATTTGCTCCCTGTCTCATGGAGA 86916

RESULT 11
BC003617
LOCUS
DEFINITION Homo sapiens RAB6A, member RAS oncogene family, transcript variant
1, mRNA (cdna clone MGC:1654 IMAGE:3506585), complete cds.
ACCESSION BC003617
```

```
BC003617.2 GI:33870607
MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3070)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.E., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
REFERENCE 2 (bases 1 to 3070)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NTH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT On Aug 19, 2003 this sequence version replaced gi:13177663.
Contact: MGC help desk
Email: cgaps-x@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 7 Row: a Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19923230.
FEATURES
Location/Qualifiers
1..3070
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:1654 IMAGE:3506585"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH MGC_14"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1..3070
/genes="RAB6A"
/notes="synonym: RAB6"
/db_xref="LocusID:5870"
/db_xref="MIM:179513"
418..1044
/codon_start=1
/product="RAB6A, member RAS oncogene family, isoform a"
source
gene
CDS
```

/protein id="AAH03617.1"
/db_xref="GI:13177664"
/db_xref="LocusID:5870"
/db_xref="MIM:179513"
translation="MSTGDFGNLRKFLVFLGQSVKTSILTRFMVDSFDNTYQA
TIGIDPLSKTYLEDTIRLQWDTAGQERFSLPSYIRDSAAVYVDITVNSFQ
QTTKWIDVTRGSDVILMLVGNKTDLADKQVSIIEGKERKAKELNMFIBTSKAG
YNNVKQLFRVAAALPGMESTQDRSDMIDIKLEKPEQPVSEGGCSC"

ORIGIN

Alignment Scores:
Pred. No.: 5, 9e-102 Length: 3070
Score: 1147.00 Matches: 228
Percent Similarity: 90.70% Conservative: 6
Best Local Similarity: 88.37% Mismatches: 20
Query Match: 87.96% Indels: 4
DB: 9 Gaps: 1

US-09-441-857-12 (1-254) x BC003617 (1-3070)

Qy 1 MetSerAlaGlyGlyAspPheGlyAenProLeuArgLysPheLysLeuValPheLeuGly 20
|||||
Db 418 ATGTCCACGGCGGAGACTTCGGGAATCCGCTGAGGAATTCAGCTGGTGTCTCTGGG 477
Qy 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAspAen 40
|||||
Db 478 GAGCAAGCGTTGGAAGACATCTTTGATCACCAGATTTCATGTATGACAGTTTGACAC 537
Qy 41 ThrTyrGlnAlaIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
|||||
Db 538 ACCTATCAGCAACAATTCGCTATTCATTTTATCAAAACATATGTACTTGGAGGATCGA 597
Qy 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgSerLeuLeuPro 80
|||||
Db 598 ACAATCAGCTTCGCTGTGGATCTGCGGTGAGGAACGTTTCGTAGCTCATTC 657
Qy 81 ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAenSer 100
|||||
Db 658 AGTTACATCCGTTCTCTGCTGAGCTGTAGTATGTTACGATATCACAATGTGTAAC 717
Qy 101 PheGlnGlnThrThrLysTrpIleAspValArgThrGluArgLysSerAspValle 120
|||||
Db 718 TTCCAGCAAACTACAAAGTGGATTCATGATGTGCAGAACAGAAAGGAGTGTATC 777
Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGlu 140
|||||
Db 778 ATCATGCTAGTAGGAATATAACAGATCTTGTCTGACAGAGCAAGTGTCAATTGAG 837
Qy 141 GlyLysArgLysAlaLysGlyLeuAenValThrPheIleGluThrArgAlaLysThrGly 160
|||||
Db 838 CGAGAGAGCAAGCAAGCAAGCTGAATGTTATGTTTATTGAAACTAGTCAAAAGCTGA 897
Qy 161 TyrAenValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr 180
|||||
Db 898 TACAATGTAAACAGAGCTCTTCGAGTGTAGCAGCAGCTTTCCCGGAATGGAAGCACA 957
Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluThr 200
|||||
Db 958 CAGGACAGAGCAGAGAGATATGATTGACATAAAACTGGAAGCCTCAGGAGCAACCA 1017
Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerThrLeuProGlnLys 220
|||||
Db 1018 GTCAGTGAAGGAGGCTGTTCTCTGCTAATCTCCCATGTCTTCAACCTTCTCAGAAG 1077
Qy 221 -----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAenLeu 236
|||||
Db 1078 TCACTGCTTTGGCCCCCTTACTCTTTTCATGTAGTGCAGTGTGAATATTCGCTTGAAC 1137
Qy 237 PheProSerLeuIleThrPheCysAenSerSerLeuLeuProValSerTrpArg 254
|||||
Db 1138 TTCCCTTCAGTAATAACGATTATGCAATTCATCTGCTCCCTGCTCTCGTGGAGA 1191

RESULT 12

BC068486
LOCUS

DEFINITION
Homo sapiens RAB6A, member RAS oncogene family, mRNA (cDNA clone

MG:87150 IMAGE:30342562), complete cds.
ACCESSION
BC068486

VERSION
BC068486.1 GI:46249770
KEYWORDS
MGC.

SOURCE
Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 3100)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Heintz, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 3100)
Strausberg, R.

Direct Submission
Submitted (02-APR-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mhc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Stefan Hansson
DNA Library Preparation: Michael Brownstein / Ted Ustin
Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 168 Row: 1 Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19923230.

Location/Qualifiers
1. 3100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87150 IMAGE:30342562"
/tissue_type="Placenta, normal"
/clone_lib="NIH MGC_147"
/lab_host="DH10B"
/note="Vector: pBluescript"

1. 3100
/gene="RAB6A"
/note="synonym: RAB6"
/db_xref="LocusID:5870"

FEATURES
source

gene


```

||||| 914 GGAGAGAGAAAGCCAAAGAGCTGAATGTATGTTTATTGAACTAGTGCAAGAGCTGA 973
|||||
Qy 161 TyrAsnValLysGlnLeuPheArgVallAlaLeuProGlyMetGluSerThr 180
|||||
Db 974 TACAATGTAAGACAGCTCTTTTCGAGCTGTAGCAGCAGCTTTGCCGGGAATGGAAGACACA 1033
|||||
Qy 181 GlnAspGlySerArgGluAsnMetSerAspIleLysLeuGluLysProGlnGlnThr 200
|||||
Db 1034 CAGGACAGAGAGAGAGATATGATTGACATAAACTGGAAGAACCCCTCAGAGAACCA 1093
|||||
Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
|||||
Db 1094 GTACGTGAGGAGGCTGTTCTCGTAAATCTCCCATGTCTATCTTCAACCTTCTTCAGAGC 1153
|||||
Qy 221 -----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
|||||
Db 1154 TCACGTGCTTTGGCCCCCTTACTCTTTCATTGACTGACGTGTGAATATTTGGCTTGAACCTT 1213
|||||
Qy 237 PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
|||||
Db 1214 TTCCCTTCAGTAATAACGTAATGCAATTCATCATTTGCTGCTGTCTCGTGAGA 1267
|||||

```

RESULT 15

```

BC019118
LOCUS BC019118 1770 bp mRNA linear ROD 04-OCT-2003
DEFINITION Mus musculus RAB6, member RAS oncogene family, mRNA (cDNA clone
MGC:29434 IMAGE:3711087), complete cds.
ACCESSION BC019118
VERSION BC019118.1 GI:17512289
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1770)
Straubeberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altechul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Sheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

```

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

TITLE

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

22388257

REFERENCE

2 (bases 1 to 1770)

AUTHORS

Straubeberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (07-DEC-2001)

National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 39 Row: a Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: GenomeScan gene
prediction.

FEATURES

source

```

1..1770
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:29434 IMAGE:3711087"
/tissue_type="Mammary tumor: C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

```

gene

```

1..1770
/gene="Rab6"
/db_xref="LocusID:19346"
/db_xref="MGI:894313"
271..897
/gene="Rab6"
/codon_start=1
/product="Rab6 protein"
/protein_id="AAH19118.1"
/db_xref="GI:17512290"
/db_xref="LocusID:19346"
/db_xref="MGI:894313"

```

CDS

```

1..1770
/translation="MSAGGDFGNPLKFKLVFGEOSVGKTSILTRMYDSEFNTYOA
TIGIDPLSKTWLYEDTIRQLQNDTAGOERFSLIYSYIRDSAAVVVYDITVNSFQ
QTTKWIDDDVRVTEGSDVIIMLVGNKTDLDARQVSIIEGGERKAKELNMVFIETSAKAG
YNYVKQLFRVRAALPGMESTQDRSREDMIDIKLEKPEQPVNEGSC"
310..792
/gene="Rab6"
/note="RAB: Region: Rab subfamily of small GTPases"
/db_xref="CDD:smart00175"

```

misc_feature

```

2..71e-101
Length: 1770
Matches: 227
Score: 1137.00
Percent Similarity: 90.31%
Conservative: 6
Best Local Similarity: 87.98%
Mismatch: 21
Query Match: 4
DB: 10
Gaps: 1

```

ORIGIN

Alignment Scores:

```

Pred. No.: 2..71e-101
Score: 1137.00
Percent Similarity: 90.31%
Conservative: 6
Best Local Similarity: 87.98%
Mismatch: 21
Query Match: 4
DB: 10
Gaps: 1

```

US-09-441-857-12 (1-254) x BC019118 (1-1770)

Qy 1 MesSerLagGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20

Db 271 ATGTCCCGGGCGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGTGTTCTCTCGGA 330

Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAsn 40

Db 331 GAGCAGAGCGCTTGGAAAGACGTCCTTGATCACCCTGATTCATGATGACGTTTTCACAC 390

Qy 41 ThrTyrGlnAlaIlelleGlylleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60

```

Db 391 ACCTATCAGGCAACAATTGGCATTGACCTTTATCAAAAACAATGTACTTGGAGGATAGA 450
QY 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 451 ACCATCAGGCTCAGCTGGGATACTCGGGTCAGGAACGTTTCGTAGCCTCATTTCCC 510
QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
Db 511 AGTTACATCCGTGACTCTGCTCAGCGGTAGTAGTTTACGATATCACAAATGTTAACTCA 570
QY 101 PheGlnGlnThrThrIleAspValAlaArgThrGluArgGlySerAspValIle 120
Db 571 TTCCAGCAAACTACAAAATGGATTGATGATGTGAGAACAGAGAGAGAGAGTGTATCATC 630
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 631 ATCATGTAGTAGGAATAAACAAGATCTTGTGATAGAGGCAAGTGTCAATTGAGGAG 690
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 691 GGAGAGAGGAAAGCCAAAGAGCTGAATGTTATGTTTATTGAAACCAGTGCAAAAGCAGGA 750
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 751 TACAATGTAAAGCAGCTTTCCGGCGTGTGTGTCAGCTTTACCTGGAATGGAAGCACA 810
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThr 200
Db 811 CAGGACAGAAAGCAGAGACATGATTGACATAAAACTGGAAGAGCCTCAGGAGCAACCA 870
QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 871 GTCAATGAGGCGGCTGTTCTGCTAAATCCCTCGCATCTTTCAACCCCTCTGCAGAG 930
QY 221 -----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
Db 931 CTCACCTGCTTTGGCCCCCATACTCTTTCATTGACTGCAGTGTGAATATTGGCTTGAACCTT 990
QY 237 PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 991 TTCCCTTCAGTAATAACGATTGCAATTCATTCCTGCTGCTGCTTTCGTGGAGG 1044

```

Search completed: April 25, 2005, 04:55:04
 Job time : 3685 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2005, 22:09:40 ; Search time 452 Seconds

(without alignments)
3326.580 Million cell updates/sec

Title: US-09-441-857-12

Perfect score: 1304

Sequence: 1 MSAGDFGPNLKRKFLVFLG.....NLFFSLITFCNSSLIPVSWR 254

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues.

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q=/Cgn2_1/USPTO_spool_p/US0441857/runat_22042005_155508_24737/app_query.fasta_1.391
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09441857 @Cgn_1_1.708 @runat_22042005_155508_24737 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1304	100.0	762	3	AAZ93839
2	1304	100.0	3084	5	ABX71303 Human int
C 3	1299	99.6	349981	10	ADC87619 Human GPC
C 4	1299	99.6	349989	10	ADC86916 Human GPC
5	1147	88.0	1460	3	AAZ93840

6	1147	88.0	3079	12	ADN03693	Adn03693 Antipori
7	1143	87.7	719	3	AAZ93834	AAZ93834 WTH3 5' p
8	1138	87.3	3745	12	ADL12658	Adl12658 Human ste
9	1133	86.9	3118	3	AAF15989	Aaf15989 Human pro
10	982	75.3	739	3	AAZ93836	AAZ93836 RAB6C cod
11	959	73.5	718	3	AAZ93835	AAZ93835 RAB6 codi
12	959	73.5	740	3	AAZ93837	AAZ93837 RAB6 codi
13	953	73.1	624	3	AAZ93841	AAZ93841 RAB6C cod
14	892	68.4	1219	4	AAK52711	AAK52711 Human pol
15	892	68.4	1266	12	ADP07320	Adp07320 Human RAB
16	892	68.4	1266	13	ADR25795	ADR25795 Breast ca
17	892	68.4	1461	4	AAK51727	AAK51727 Human pol
18	884	67.8	1589	2	AAZ11736	AAZ11736 Human tra
19	883	67.7	573	3	AAZ93838	AAZ93838 WTH3 codi
20	865.5	66.4	3195	5	AAZ93832	AAZ93832 DNA encod
21	846.5	64.9	1188	4	ABL11393	Abi111393 Drosophil
22	846.5	64.9	1188	4	ABL11392	Abi111392 Drosophil
23	835.5	64.1	2456	10	ADA52862	Ada52862 Human cod
24	715	54.8	926	3	AAF14485	Aaf14485 Aspergill
25	712	54.6	1074	3	AAC33337	Aac33337 Arabidops
26	709.5	54.4	1355	3	AAC46575	Aac46575 Zea may
27	673.5	51.6	624	3	AAC42719	Aac42719 Arabidops
28	667	51.2	523	13	ACN59588	Acn59588 Cotton gy
29	665	51.0	636	10	ADC76216	Adc76216 DNA homol
30	665	51.0	636	10	ADK55884	Adk55884 Plant DNA
31	643	49.3	645	3	AAC43018	Aac43018 Arabidops
32	642	49.2	443	8	ABX42190	Abx42190 Bovine ES
33	627	48.1	646	10	ADD16517	Add16517 DNA (Seqi
34	627	48.1	646	11	ADM44783	Adm44783 Insect re
35	627	48.1	647	10	ADC75668	Adc75668 DNA homol
36	627	48.1	723	10	ADC76194	Adc76194 DNA homol
37	627	48.1	723	10	ADK58490	Adk58490 Plant DNA
38	612.5	47.0	495	3	AAC41466	Aac41466 Zea may
39	596	45.7	791	13	ADR65699	Adr65699 Cotton cd
40	596	45.7	928	6	ABI99630	Abi99630 Mouse iac
41	593.5	45.5	3826	10	ADC86930	Adc86930 Human GPC
42	586.5	45.0	642	5	AAZ93839	AAZ93839 Human int
43	582	44.6	648	13	ADA47056	Ada47056 Bacterial
44	572	43.9	846	13	ADS47382	Ads47382 Bacterial
45	569	43.6	424	9	ACH18328	Ach18328 Human adu

ALIGNMENTS

RESULT 1
AAZ93839
ID AAZ93839 standard; DNA; 762 BP.

AC AAZ93839;

DT 29-AUG-2000 (first entry)

DE WTH3 coding sequence.

XX WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;
KW antibody; immunogen; mutation; detection; therapy; human; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..762

FT FT /*tag= a

FT FT /product= "WTH3"

XX PN WO200029625-A1.

XX PD 25-MAY-2000.

XX PF 18-NOV-1999; 99WO-US027630.

XX PR 18-NOV-1998; 98US-0108994P.

XX PA (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.

Db 45279 GGAGAGAGAAAGCCAAAGGGCTGAATGTACGTTTATTGAACATAGGCGCAAAAGCTGGA 45220
Qy 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 45219 TACAATGTAAAGCAGCTCTTTGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAGACACA 45160
Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 45159 CAGGACGGAAGCAGAGAGACATGAGTGACATAAACTGGAAAGCCTCAGAGGCAACA 45100
Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 45099 GTCAGCGAAGGGGGTGTCTCTGCTACTCTCCCATGTCTATCTCAACCCCTTCTCAGAAG 45040
Qy 221 ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeu 240
Db 45039 CCCCCCTTACTCTTTCAATGACGTCAGTGTGAATATTTGGCTTGAACCTTTTCCCTTCATTA 44980
Qy 241 IleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 44979 ATAACGTTTTGCAATTCATCATTTGCTGCTGTCTGTGGAGG 44938

RESULT 4
ADC86916/c
ID ADC86916 standard; DNA; 349989 BP.
XX AC ADC86916;
XX ds; gene; human; GPCR;
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX Homo sapiens.
XX EP1270724-A2.
XX 02-JAN-2003.
XX 18-JUN-2002; 2002EP-00013517.
XX 18-JUN-2001; 2001JP-00246789.
XX (NIAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX Suwa M, Asai K, Akiyama Y, Aburatani H;
XX WPI; 2003-315783/31.
DR P-FSDB; ADC86917.

New polynucleotide, useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor.
PS Claim 1; SEQ ID NO 1369; 28pp; English.

XX The invention relates to a novel polynucleotide encoding a guanosine triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of the invention may have a use in gene therapy. The polynucleotide and polypeptide are useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor. The polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the invention.
SQ Sequence 349989 BP; 93096 A; 80097 C; 83967 G; 92327 T; 0 U; 502 Other;

Alignment Scores:
Pred. No.: 5.51e-135 Length: 349989
Score: 1299.00 Matches: 253

Percent Similarity: 99.61% Conservative: 0
Best Local Similarity: 99.61% Mismatches: 1
Query Match: 99.62% Indels: 0
DB: 10 Gaps: 0
US-09-441-857-12 (1-254) x ADC86916 (1-349989)
Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 345708 ATGTCGCGGCGGAGACTTCGGGAATCCGCTGAGGAATTCAGAGCTGGTGTCTCTGGGG 345649
Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40
Db 345648 GAGCAAAAGCGTTGCAAAAGACATCTTTGATCACCAGATTCAGGTATGACAGTTTTTGACAAC 345589
Qy 41 ThrTyrGlnAlaIleIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 345588 ACCTATCAGGCAATAATTTGGCATTGACTTTTTTATCAAAAACATATGTACTTTGGAGGATGGA 345529
Qy 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 345528 ACAATCGGGCTTCGGCTGTGGGATACGGGGTTCAGGAACGTCTCCGTAGCCTCATTTCCC 345469
Qy 81 ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAsnSer 100
Db 345468 AGGTACATCCGTGATTCTGCTGCAGCTGTAGTAGTTTACGATATCACAAATGTTAACTCA 345409
Qy 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 345408 TTCCAGCAAACTACAAAAGTGGATTGATGATGTCAAGACAGAAAGAGGAGTGATGTATC 345349
Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 345348 ATCAGCTAGTAGGAATAGAACAGATCTTCTGCACAGAGGCAAGTGTCTAGTTGAGAG 345289
Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 345288 GGAGAGAGAAAGCCAAAGGGCTGAATGTTTACGTTTATTGAAACTAGSGCAAAAGCTGGA 345229
Qy 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLysLeuProGlyMetGluSerThr 180
Db 345228 TACAATGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAGACACA 345169
Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 345168 CAGGACGGAAGCAGAGAAACATGATGATGACATAAACTGGAAAGCCTCAGAGGCAACA 345109
Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 345108 GTCAGCGAAGGGGGTGTCTCTGCTACTCTCCCATGTCACTTCAACCCCTTCTCTCAGAAG 345049
Qy 221 ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeu 240
Db 345048 CCCCCCTTACTCTTTCAATGACGTCAGTGTGAATATTTGGCTTGAACCTTTTCCCTTCATTA 344989
Qy 241 IleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 344988 ATAACGTTTTGCAATTCATCATTTGCTGCTGTCTGTGGAGG 344947
RESULT 5
AAZ93840
ID AAZ93840 standard; DNA; 1460 BP.
XX AC AAZ93840;
XX 29-AUG-2000 (first entry)
DT 29-AUG-2000 (first entry)
XX RAB6C coding sequence.
XX WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;
KW antibody; immunogen; mutation; detection; therapy; human; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 420..1046
 FT /*tag= a
 FT /product= "RAB6C"
 XX WO200029625-A1.
 PN 25-MAY-2000.
 PD 18-NOV-1999; 99WO-US027630.
 PP 18-NOV-1998; 98US-0108994P.
 XX (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.
 XX Duffy HX, Shan J, Yuan L, Budman D, Calabro A;
 XX WPI; 2000-387828/33.
 DR P-PSDB; AAY83403.
 XX Novel nucleic acids encoding wt3 protein useful for inhibiting multiple
 PT drug resistance in tumor cells for identifying the modulators of drug
 PT resistance.
 XX Claim 6; Page 68-70; 82pp; English.
 XX Nucleic acids encoding WTH3 and RAB6 were isolated by Methy1- (or mutant)
 CC Differential Display (MDD) methods and can be used in methods for
 CC detecting methylation patterns in multiple drug resistance loci. Genes
 CC are frequently not methylated in cells where they are expressed but are
 CC methylated in cell types where they are not expressed. Tumour cell DNA is
 CC often methylated to a different extent and in different regions when
 CC compared to DNA of normal cells. The methylation pattern in a multiple
 CC drug resistance locus can be altered and give rise to altered expression
 CC patterns of that multiple drug resistance locus. Nucleic acids
 CC corresponding to the identified loci such as WTH3 and RAB6 nucleic acids
 CC can be used as probes for detecting mutations and methylation patterns of
 CC those loci. The nucleic acids and their homologues are useful for
 CC inhibition of multiple drug resistance and for treating tumors exhibiting
 CC multiple drug resistance. They are also useful for detecting and
 CC measuring the expression of mRNA from identified genes and for
 CC determining suitable therapeutic treatment. Antibodies directed against
 CC immunogenic fragments of WTH3 and RAB6 are useful for detecting specific
 CC proteins and polypeptides in tissues or body fluids of patients
 XX
 SQ Sequence 1460 BP; 365 A; 356 C; 358 G; 381 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5,06e-121 Length: 1460
 Score: 1147.00 Matches: 228
 Percent Similarity: 90.70% Conservative: 6
 Best Local Similarity: 88.37% Mismatches: 20
 Query Match: 87.96% Indels: 4
 DB: 3 Gaps: 1

US-09-441-857-12 (1-254) x AA293840 (1-1460)

QY 1 MetSerAlaGlyClyAspPheGlyAanProLeuArgLysPheLysLeuValPheLeuGly 20
 Db 420 ATGTCCACCGCGGAGACTTCGGGAATCCGCTGAGGAAATTCGAAGCTGTGCTCGGG 479
 QY 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAan 40
 Db 480 GAGCAAGCGTTGGGAAGACATCTTGATCACCAGATTCATGTATGACAGTTTGGACAC 539
 QY 41 ThrTyrGlnAlaIleGlyLeuPheLeuSerLysThrMetTyrLeuGluAspGly 60
 Db 540 ACCATCAGCAACAATTGGCATTGACTTTTATCAAAACATGTACTTGGAGGATCGA 599
 QY 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuSerLeuIlePro 80
 Db 600 ACAATCAGGCTTCAGCTGTGGGATACTGCGGGTCAGGAACGTTTCGTCAGCCTCATCCC 659

QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
 Db 660 AGTTACATCCGTGATTCTGCTGCAGCTGTAGTAGTTTACGATATACAAATGTTAACTCA 719
 QY 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgLysSerAspValIle 120
 Db 720 TTCCAGCAAACTACAAAGTGGATTGATGTGTGACAGACAGAAAGAGGAGTGTATGTC 779
 QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
 Db 780 ATCATGCTAGTAGGAATATAACAGATCTTGTCTGCAAGAGCAAGTGTCAATTGAGGAG 839
 QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
 Db 840 GGAGAGAGGAAAGCCAAAGAGCTGAATGTTATGTTTATTGAAACTAGTGCNAAAGCTGA 899
 QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
 Db 900 TACAATGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAAGCACA 959
 QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
 Db 960 CAGGACAGAGCAGAGAGATATGATTGACATATAAACTGGAAAGCCTCAGGAGCAACCA 1019
 QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
 Db 1020 GTCACTGAAGGAGGCTGTTCTCTGCTTAATCTCCCATGTCTCATCTTCAACCTTCTTCAGAAGC 1079
 QY 221 -----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
 Db 1080 TCACCTGCTTTGGCCCCCTTACTCTTTTCATTGACTGCAGTGTGAATATTGGCTTGAACCTT 1139
 QY 237 PheProSerLeuLeuThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
 Db 1140 TTCCCTTCAGTAATAACGTAATGCAATTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1193

RESULT 6
 ADN03693
 ID ADN03693 standard; cDNA; 3079 BP.
 XX AC ADN03693;
 XX DT 01-JUL-2004 (first entry)
 XX DE Antipsoriatic cDNA sequence #44.
 XX KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
 XX OS Homo sapiens.
 XX PN WO2004028479-A2.
 XX PD 08-APR-2004.
 XX PF 25-SEP-2003; 2003WO-US030907.
 XX PR 25-SEP-2002; 2002US-0414006P.
 XX PA (GETH) GENENTECH INC.
 XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 XX Wu TD;
 XX WPI; 2004-305105/28.
 XX P-PSDB; ADN03694.
 XX New PRO nucleic acid or polypeptide, useful for preparing a
 XX pharmaceutical composition for diagnosing or treating psoriasis in a
 XX mammal.
 XX Claim 1; SEQ ID NO 87; 3069pp; English.

The invention relates to novel polynucleotide and polypeptides for treating psoriasis or a sequence having at least 80% identity to the above sequences. The nucleic acid is useful for preparing a composition for diagnosing or treating psoriasis in a mammal. This sequence corresponds to one of the polynucleotides of the invention.

Alignment Scores:		
Pred. No.:	1.49e-120	Length:
Score:	1147.00	Matches:
Percent Similarity:	90.70%	Conservative:
Best Local Similarity:	88.37%	Mismatches:
Query Match:	87.96%	Indels:
DB:	12	Gaps:
		3079
		228
		6
		20
		4
		1

US-09-441-857-12 (1-254) x ADN03693 (1-3079)

Qy	1	MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly	20
Db	427	ATGTCCACGGCGGAGACTTCGGGAATCCGTGACGAAATTCAGACTGGTCTCTGGGG	486
Qy	21	GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn	40
Db	487	GAGCAAAAGCGTTGGAAAGACATCTTTGATCACCAGATTCATGTATGACAGTTTTGACAA	546
Qy	41	ThrTyrGlnAlaIleIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly	60
Db	547	ACCTATCAGGCAACAAATGGCATTTGACCTTTTATCAAAAACATAATGACTCTTGAGGATCGA	606
Qy	61	ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro	80
Db	607	ACAATCAGGCTTCAGCTGTGGATCTCGGGCTCAGGAACGTTCCGTAGGCTCATTTCCC	666
Qy	81	ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAsnSer	100
Db	667	AGTTACATCCGTGATTCGCTGCAGCTGTAGTGTTCAGGATATCACAAATGTTAACTCA	726
Qy	101	PheGlnGlnThrThrLysTyrIleAspAspValArgThrGluArgGlySerAspValIle	120
Db	727	TTCCAGCAAACTACAAAGTGATTCATGATGTCAAGACAGAAAGAGGAAGTGAATGTTATC	786
Qy	121	IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu	140
Db	787	ATCATGTGTAGTAGAAATATAAACAGATCTTGCTCACAAGAGCGCAAGTGTCAATTCAGGAG	846
Qy	141	GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly	160
Db	847	GGAGAGGAGAAAGCCAAAGAGCTGAATGTTATGTTTATGAAACTAGTGCAAAAGCTCGA	906
Qy	161	TyrAsnValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr	180
Db	907	TACAAATGTAAAGCAGCTCTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAAGCAC	966
Qy	181	GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluInThr	200
Db	967	CAGGACAGAAGCAGAGAAGATATGATTGACATAAACTGGAAAAGCCCTCAGGAGCAACCA	1026
Qy	201	ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys	220
Db	1027	GTCAGTGAAGAGAGCTGTCTCTGCTAAATCTCCCATGTATCTTCAACCTTCTTCAGAAAGC	1086
Qy	221	-----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu	236
Db	1087	TCACTGCTTTGGCCCCCTTACTCTTTCATTGACTGCAGTGTGAATATTTGGCTTGAACCTT	1146
Qy	237	PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTyrArg	254
Db	1147	TTCCCTTCAGTAATAACGATTTGAATTCATTCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200

RESULT 7
AAZ93834
ID AAZ9

XX	AAZ93834;	
AC		
XX		
DT	29-AUG-2000 (first entry)	
XX		
DE	WTH3 5' partial sequence.	
XX		
KW	WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;	
KW	antibody; immunogen; mutation; detection; therapy; human; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	49..719
FT		/*tag= a
FT		/product= "WTH3 C-terminal fragment"
XX		
PN	WO200029625-A1.	
XX		
PD	25-MAY-2000.	
XX		
PF	18-NOV-1999; 99WO-US027630.	
XX		
PR	18-NOV-1998; 98US-0108994P.	
XX		
PA	(NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.	
XX		
PI	Duffy HX, Shan J, Yuan L, Budman D, Calabro A;	
XX		
DR	WPI; 2000-387828/33.	
DR	P-PSDB; AAY83398.	
XX		
PT	Novel nucleic acids encoding wth3 protein useful for inhibiting multiple	
PT	drug resistance in tumor cells for identifying the modulators of drug	
PT	resistance.	
XX		
PS	Claim 19; Fig 1a-1b; 82pp; English.	
XX		
CC	Nucleic acids encoding WTH3 and RAB6 were isolated by Methy1- (or mutant)	
CC	Differential Display (MDP) methods and can be used in methods for	
CC	detecting methylation patterns in multiple drug resistance loci. Genes	
CC	are frequently not methylated in cells where they are expressed but are	
CC	methylated in cell types where they are not expressed. Tumour cell DNA is	
CC	often methylated to a different extent and in different regions when	
CC	compared to DNA of normal cells. The methylation pattern in a multiple	
CC	drug resistance locus can be altered and give rise to altered expression	
CC	patterns of that multiple drug resistance locus. Nucleic acids	
CC	corresponding to the identified loci such as WTH3 and RAB6 nucleic acids	
CC	can be used as probes for detecting mutations and methylation patterns of	
CC	those loci. The nucleic acids and their homologues are useful for	
CC	inhibition of multiple drug resistance and for treating tumors exhibiting	
CC	multiple drug resistance. They are also useful for detecting and	
CC	measuring the expression of mRNA from identified genes and for	
CC	determining suitable therapeutic treatment. Antibodies directed against	
CC	immunogenic fragments of WTH3 and RAB6 are useful for detecting specific	
CC	proteins and polypeptides in tissues or body fluids of patients	
XX		
SQ	Sequence 719 BP; 208 A; 154 C; 193 G; 164 T; 0 U; 0 Other;	
	Alignment Scores:	
	Pred. No.:	Length: 719
	Score:	1143.00 Matches: 224
	Percent Similarity:	100.00% Conservative: 0
	Best Local Similarity:	100.00% Mismatches: 0
	Query Match:	87.65% Indels: 0
	Deby	3 Gaps: 0
	US-09-441-857-12 (1-254) x AAZ93834 (1-719)	

```
QY 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAspAen 40
Db 108 GAGCAAGCGTTCGCAAGACATCTTTGATCACCAGATTGAGTATGACAGATTTTGCACAA 167
QY 41 ThrTyrGlnAlaIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 168 ACCTATCAGGCAATTAATTGGCATTGACTTTTATCAAAACATATGTACTTGGAGATGGA 227
QY 61 ThrLeuGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuLeuPro 80
Db 228 ACAATCGGGCTTCGGCTGTGGATACGCGGGTACGGAACGCTCTCGTAGCCTCATTC 287
QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
Db 288 AGGTACATCCGTGATTCGCTGCAGCTGTAGTAGTTTACGATATCACAAATGTAACTCA 347
QY 101 PheGlnGlnThrTyrThrTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 348 TTCCAGCAAACTACAAAGTGGATTGATGATGTCAGAACAGAAAGAGGAGTGTATGTTATC 407
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 408 ATCAGCTAGTAGGAAATAGAACAGATCTTGCTGACAGAGGCAAGTGTCACTTCAGGAG 467
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 468 GGAGAGAGGAAAGCCAAAGGGCTGAATGTTACGTTTATTGAAACTAGGGCAAAACTGGA 527
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr 180
Db 528 TACAATGTAAAGCAGCTCTTTCCGCGTGTAGCAGCAGCTTTCCCGGGAATGGAAAGCACA 587
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThr 200
Db 588 CAGGACGGNAGCAGAGACAGATGAGTGACATTAACCTGGAAAGCCCTCAGGAGCAACA 647
QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 648 GTCAGCGAAGGGGTGTTCTCTGCTACTCTCCCATGTCTATCTTCAACCCCTCTCAGAAG 707
QY 221 ProProTyrSer 224
Db 708 CCCCTTACTCT 719

RESULT 8
ADL12658
ID ADL12658 standard; cDNA; 3745 BP.
AC ADL12658;
DT 06-MAY-2004 (first entry)
XX Human steroid-induced C3A liver cell cDNA #387.
DE ss; Gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX OS Homo sapiens.
XX PN US6673549-B1.
XX PD 06-JAN-2004.
XX PF 12-OCT-2001; 2001US-00976594.
XX PR 12-OCT-2000; 2000US-0240409P.
XX PA (INCY-) INCYTE CORP.
XX PI Furness LM, Buchbinder JL;
XX WPI; 2004-068610/07.
XX
```

```
PT Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX Claim 1; SEQ ID NO 387; 141pp; English.
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed from
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 3745 BP; 1066 A; 755 C; 745 G; 1179 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 2,13e-119 Length: 3745
Score: 1138.00 Matches: 225
Percent Similarity: 89.92% Conservative: 7
Best Local Similarity: 87.21% Mismatches: 22
Query Match: 87.27% Indels: 4
Db: 12 Gaps: 1

US-09-441-857-12 (1-254) x ADL12658 (1-3745)
```

```
QY 1 MetSerAlaGlyLysAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 494 ATGTCACCGGGCGGAGACTTCGGGNATCCGTCGAGAAATTCAGCTGGTGTCTCTGGGG 553
QY 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAen 40
Db 554 GAGCAAGCGTTCGGAAGACATCTTTGATCACCAGATTTCATGTATGACAGATTTTGCACAA 613
QY 41 ThrTyrGlnAlaIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 614 ACCTATCAGGCAAACTTGGCATTGACTTTTATCAAAACATATGTACTTGGAGATCGA 673
QY 61 ThrLeuGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 674 ACAGTACGATTGCAATTAATGGACACAGAGCTCAAGAGCGGTCAGAGAGCTTGATTCCT 733
QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
Db 734 AGCTACATTCGTGACTCCACTGTCAGTGTGTTGTTTATGATATACAAAATGTTAACTCA 793
QY 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 794 TTCCAGCAAACTTACAAAGTGGATTGATGATGTCAGAACAGAAAGAGGAGTGTATGTTATC 853
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 854 ATCATGCTAGTAGGAAATATAAACAGATCTTCTGCAAGAGGCAAGTGTCAATTCAGGAG 913
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 914 GGAGAGAGGAAAGCCAAAGAGCTGAATGTTATGTTTATTGAAACTAGTGTCAAAAGCTGA 973
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr 180
```

Db 974 TACATGTAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTCCCGGGAATGGAAGCACA 1033
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLeuGluLeuProGlnGlnThr 200
Db 1034 CAGGACAGAAGCAGAGAATATGATTTGACATAAACTGGAAGAGCCTCAGGAGCAACCA 1093
QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 1094 GTCACTGAAGGAGGCTGTTCCTGCTAAATCTCCCATGTCTTCAACCTTTCTTCAGAAAGC 1153
QY 221 -----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
Db 1154 TCACGTCTTTGGCCCCCTTACTCTTTTCATTGACTGTCAGTGTGAATATTGGCTTGAACCTT 1213
QY 237 PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 1214 TTCCCTTCAGTAATAACGATATTGCAATTTCATCTGCTGCCTGTCTCGTGAGA 1267

RESULT 9

ID AAF15989 standard; cDNA; 3118 BP.
AC AAF15989;
XX
DT 13-MAR-2001 (first entry)
XX
XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:424.
DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
OS Homo sapiens.

XX WO200055174-A1.
PN 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005988.
PR 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-587513/55.
DR P-PSDB; AAB56786.
XX
XX Prostate cancer associated gene sequences, referred to as prostate cancer
XX antigens, useful for treatment, prevention, and diagnosis of disorders
XX such as prostate cancer.

XX Claim 1; Page 906; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX The prostate cancer antigens can have neuroprotective, cytostatic,
XX cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX nephrotropic, antiinfective, gynaecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen.
XX polynucleotides may be used for detection of prostate cancer.
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX AAB57303 represent sequences used in the exemplification of the present
XX invention

XX SQ Sequence 3118 BP; 938 A; 636 C; 638 G; 902 T; 0 U; 4 Other;
Alignment Scores:
Pred. No.: 6.12e-119 Length: 3118
Score: 1133.00 Matches: 224
Percent Similarity: 89.53% Conservative: 7
Best Local Similarity: 86.82% Mismatches: 23
Query Match: 86.89% Indels: 4
DB: 3 Gaps: 1
US-09-441-857-12 (1-254) x AAF15989 (1-3118)
QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 415 ATGTCCACGGCGGAGACTTTCGGGAATCCGCTGAGGAAATTCAGAGCTGGTTCCTGGGG 474
QY 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40
Db 475 GAGCAAGCKNTGGAAAGACATCTTTGATCACCAGATTCATGTATGACAGTTTGTGCAAC 534
QY 41 ThrTyrGlnAlaIleIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 535 ACCTATCAGGCAACAATTGGCAITGACTTTTATCAAAACTATGTACTTGGAGATCGA 594
QY 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 595 ACAGTACGATTGCAATTTATGGGACACAGCAGGTTCAGAGCGTTTCAGAGCTTGTATCCT 654
QY 81 ArgTyrIleArgSerAlaAlaValValValValValValValValValValValValVal 100
Db 655 AGCTACATTCGTGACTCCACTGTGGCAGTTGTTTATGATATCACAATGTTTAATCTCA 714
QY 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 715 TTCAGCAAACTACAAAGTCGATGATGATGTGACAGACAGAGAGGAGTGTCAATTTGAGGAG 774
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 775 ATCATGTCTAGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTCGCGGAATGGAAGCACA 834
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 835 GGAGAGAGGAAAGCCAAAGAGCTGATGTTTATTGAACTAGTGTCAAAAGCTGGA 894
QY 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 895 TACAATGTAAAGCAGCTCTTTTCGACGTGTAGCAGCAGCTTTCGCGGAATGGAAGCACA 954
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThr 200
Db 955 CAGGACAGAAGCAGAGAAGATATGATTGACATAAACTGGAAAGCCCTCAGGAGCAACCA 1014
QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 1015 GTCAGTGAAGAGGCTGTTCCTGCTAAATCTCCCATGTCTTCAACCTTCTTCAGAAAGC 1074
QY 221 -----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
Db 1075 TCACGTCTTTGGCCCCCTTACTCTTTTCATTGACTGTCAGTGTGAATATTGGCTTGAACCTT 1134
QY 237 PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 1135 TTCCCTTCAGTAATAACGATATTGCAATTTCATCTGCTGCCTGTCTCGTGAGA 1188
RESULT 10
AAZ93836
ID AAZ93836 standard; DNA; 739 BP.
XX
XX AAZ93836;
XX
XX
DT 29-AUG-2000 (first entry)
XX

DE RAB6C coding sequence.
 XX WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;
 KW antibody; immunogen; mutation; detection; therapy; human; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 70..696
 FT /*tag= a
 FT /label= RAB6C polypeptide
 FT
 XX WO200029625-A1.
 PN
 XX
 XX 25-MAY-2000.
 PD
 XX
 XX 18-NOV-1999; 99WO-US027630.
 PF
 XX 18-NOV-1999; 98US-0108994P.
 PR
 XX (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.
 PA
 XX Duffy HX, Shan J, Yuan L, Budman D, Calabro A;
 PI
 XX WPI; 2000-387828/33.
 XX P-PSDB; AAY83400.
 DR
 XX
 XX Novel nucleic acids encoding wth3 protein useful for inhibiting multiple
 PT drug resistance in tumor cells for identifying the modulators of drug
 PT resistance.
 PT
 PS Example 1; Fig 3a-3b; 82pp; English.
 XX
 CC Nucleic acids encoding WTH3 and RAB6 were isolated by Methyl- (or mutant)
 CC Differential Display (MDD) methods and can be used in methods for
 CC detecting methylation patterns in multiple drug resistance loci. Genes
 CC are frequently not methylated in cells where they are expressed but are
 CC methylated in cell types where they are not expressed. Tumour cell DNA is
 CC often methylated to a different extent and in different regions when
 CC compared to DNA of normal cells. The methylation pattern in a multiple
 CC drug resistance locus can be altered and give rise to altered expression
 CC patterns of that multiple drug resistance locus. Nucleic acids
 CC corresponding to the identified loci such as WTH3 and RAB6 nucleic acids
 CC can be used as probes for detecting mutations and methylation patterns of
 CC those loci. The nucleic acids and their homologues are useful for
 CC inhibition of multiple drug resistance and for treating tumors exhibiting
 CC multiple drug resistance. They are also useful for detecting and
 CC measuring the expression of mRNA from identified genes and for
 CC determining suitable therapeutic treatment. Antibodies directed against
 CC immunogenic fragments of WTH3 and RAB6 are useful for detecting specific
 CC proteins and polypeptides in tissues or body fluids of patients
 XX
 SQ Sequence 739 BP; 219 A; 151 C; 191 G; 178 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,52e-102 Length: 739
 Score: 982.00 Matches: 196
 Percent Similarity: 92.13% Conservatives: 3
 Best Local Similarity: 90.74% Mismatches: 17
 Query Match: 75.31% Indels: 0
 DB: 3 Gaps: 0
 US-09-441-857-12 (1-254) x AA293836 (1-739)
 QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
 DB 70 ATGTCACGGCGGAGACTTCGGGAATCCGTCAGGAAATTCAGCTGGTGTCTCTGGGG 129
 QY 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAspAsn 40
 DB 130 GAGCAAGCGTTGGAAAGACATCTTTGATCACCAGATTCATGTGACAGTTTTCACAC 189
 QY 41 ThrTyrGlnAlaIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
 DB 190 ACCTATCAGGCAACAATTGGCAATTGACTTTTATCAAAAACTATGTACTTCGAGGATCGA 249
 QY 61 ThrIleGlyLeuArgLeuTtpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
 DB 250 ACAATCAGGCTTCAGCTGTGGGATCTCGGGTCAGGAACGTTTCGCTAGGCTCATTCCTC 309
 QY 81 ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAsnSer 100
 DB 310 AGTTACATCCGCTGATTCTGCTGCAGCTGTAGTAGTATACGATATACAAATGTTAACTCA 369
 QY 101 PheGlnGlnThrThrLysTtpIleAspAspValArgThrGluArgGlySerAspValIle 120
 DB 370 TTCAGCAAACTACAAAGTGGATTGATGATGTCAAGACAGAAAGAGGAAGTGTATTATC 429
 QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
 DB 430 ATCATGCTAGTAGGAATATAACAGATCTTGCTGACCAAGAGGCAAGTGTCAATTGAGGAG 489
 QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
 DB 490 GGAGAGAGGAAAGCCAAAGAGCTCAATGTTATGTTTATTGAAACTAGTGCAGAAAGCTGGA 549
 QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLalaLeuProGlyMetGluSerThr 180
 DB 550 TACAATGTAAAGCAGCTCTTTTCGACGCTGTAGCAGCAGCTTTGCCGCGGAATGGAAGCACA 609
 QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
 DB 610 CAGACAGAGACAGAGAGATATGATTGACATAAACTGGANAAGCCCTCAGGAGCAACCA 669
 QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThr 216
 DB 670 GTCAGTGAAGGAGGCTGTCTCTGCTTAATCTCCCATGTCATCTTCAACC 717
 RESULT 11
 AA293835
 ID AA293835 standard; DNA; 718 BP.
 AC AA293835;
 XX 29-AUG-2000 (first entry)
 DE RAB6 coding sequence.
 KW WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;
 KW antibody; immunogen; mutation; detection; therapy; human; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 49..675
 FT /*tag= a
 FT /product= "RAB6"
 FT
 XX WO200029625-A1.
 PN
 XX 25-MAY-2000.
 PD
 XX 18-NOV-1999; 99WO-US027630.
 PF
 XX 18-NOV-1999; 98US-0108994P.
 PR
 XX (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.
 PA
 XX Duffy HX, Shan J, Yuan L, Budman D, Calabro A;
 PI
 XX WPI; 2000-387828/33.
 XX P-PSDB; AAY83399.
 DR
 XX Novel nucleic acids encoding wth3 protein useful for inhibiting multiple
 PT drug resistance in tumor cells for identifying the modulators of drug
 PT resistance.

XX

PS Example 1; Fig 1a-1b; 82pp; English.

CC Nucleic acids encoding WTH3 and RAB6 were isolated by Methyl- (or mutant)
CC Differential Display (MDD) methods and can be used in methods for
CC detecting methylation patterns in multiple drug resistance loci. Genes
CC are frequently not methylated in cells where they are expressed but are
CC methylated in cell types where they are not expressed. Tumour cell DNA is
CC often methylated to a different extent and in different regions when
CC compared to DNA of normal cells. The methylation pattern in a multiple
CC drug resistance locus can be altered and give rise to altered expression
CC patterns of that multiple drug resistance locus. Nucleic acids
CC corresponding to the identified loci such as WTH3 and RAB6 nucleic acids
CC can be used as probes for detecting mutations and methylation patterns of
CC those loci. The nucleic acids and their homologues are useful for
CC inhibition of multiple drug resistance and for treating tumors exhibiting
CC multiple drug resistance. They are also useful for detecting and
CC measuring the expression of mRNA from identified genes and for
CC determining suitable therapeutic treatment. Antibodies directed against
CC immunogenic fragments of WTH3 and RAB6 are useful for detecting specific
CC proteins and polypeptides in tissues or body fluids of patients
XX

SQ Sequence 718 BP; 219 A; 141 C; 184 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,31e-100 Length: 718
Score: 959.00 Matches: 191
Percent Similarity: 90.28% Conservative: 4
Best Local Similarity: 88.43% Mismatches: 21
Query Match: 73.54% Indels: 0
DB: 3 Gaps: 0

US-09-441-857-12 (1-254) x AA293835 (1-718)

QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
DB 49 ATGTCACGGCGGAGACTTCGGGAATCCGCTGAGGAATTCAGCTGGTGTCTCTGGG 108

QY 21 GluGlnSerValAlaLysThrSerLeulleThrArgPheArgThrAspSerPheAspAsn 40
DB 109 GAGCAAAACGTTGGAACACATCTTTGATCACCAGATTTCATGATGACAGTTTTCACAA 168

QY 41 ThrTyrGlnAlaIleGlyLeaPheLeuSerLysThrMetTyrLeuGluAspGly 60
DB 169 ACCTATCAGGACACAAATTCGATTCATTTTATCAAAACTATGCTCTGGAGATCGA 228

QY 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
DB 229 ACAGTACGATTGCAATTATGGACACACAGCAGTCAAGAGCGTTTCAGGAGCTTGATTCCT 288

QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
DB 289 AGCTACATTCGTGACTCCACTGTCGAGTGTGTTTATGATATATACAAATGTTAACTCA 348

QY 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
DB 349 TTCAGCAAACTCAAAAGTGGATGATGATGTCAGAACAGAAAGAGGAAGTGTATC 408

QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValIleGlu 140
DB 409 ATCATGCTAGTAGGAATAAAACAGATCTTCCTGACAAAGAGCAAGTGTCAATTGAGGAG 468

QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
DB 469 GGAGAGAGGAAGCAAGAGCTGAATGTTATGTTTATTTGAACCTAGTGTCAAAAGCTGGA 528

QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
DB 529 TACAATGTAAAGCAGCTCTTTTCAGCTGTACAGCAGCTTTGCCGGAATGGAAGACACA 588

QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluInThr 200
DB 589 CAGCACAGAGCAGAGAAGATATGATTGACATATAAACTGGAAAGCCTCAGGACCAACCA 648

QY

201 ValSerGluGlyCysSerCysTyrSerProMetSerSerSerThr 216
DB 649 GTCAGTGAAGAGGCTGTTCTGCTAATGTGCTTCCCTAGTCACTTCACACC 696

RESULT 12

AA293837
ID AA293837 standard; DNA; 740 BP.

AC AA293837;

XX 29-AUG-2000 (first entry)

DE RAB6 coding sequence.

XX WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;
KW antibody; immunogen; mutation; detection; therapy; human; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 71..697
FT /*tag= a
FT /product= "RAB6"

XX WO200029625-A1.

XX 25-MAY-2000.

XX 18-NOV-1999; 99WO-US027630.

XX 18-NOV-1998; 98US-0108994P.

XX (NSHO-) NORTH SHORE-LONG ISLAND JEWISH RES.

XX Duffy HX, Shan J, Yuan L, Budman D, Calabro A;

XX WPI; 2000-387828/33.

XX P-PSDB; AAY83399.

XX Novel nucleic acids encoding wth3 protein useful for inhibiting multiple
XX drug resistance in tumor cells for identifying the modulators of drug
XX resistance.

XX Example 1; Fig 3a-3b; 82pp; English.

XX Nucleic acids encoding WTH3 and RAB6 were isolated by Methyl- (or mutant)
XX Differential Display (MDD) methods and can be used in methods for
XX detecting methylation patterns in multiple drug resistance loci. Genes
XX are frequently not methylated in cells where they are expressed but are
XX methylated in cell types where they are not expressed. Tumour cell DNA is
XX often methylated to a different extent and in different regions when
XX compared to DNA of normal cells. The methylation pattern in a multiple
XX drug resistance locus can be altered and give rise to altered expression
XX patterns of that multiple drug resistance locus. Nucleic acids
XX corresponding to the identified loci such as WTH3 and RAB6 nucleic acids
XX can be used as probes for detecting mutations and methylation patterns of
XX those loci. The nucleic acids and their homologues are useful for
XX inhibition of multiple drug resistance and for treating tumors exhibiting
XX multiple drug resistance. They are also useful for detecting and
XX measuring the expression of mRNA from identified genes and for
XX determining suitable therapeutic treatment. Antibodies directed against
XX immunogenic fragments of WTH3 and RAB6 are useful for detecting specific
XX proteins and polypeptides in tissues or body fluids of patients
XX

SQ Sequence 740 BP; 223 A; 146 C; 193 G; 178 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,59e-100 Length: 740
Score: 959.00 Matches: 191
Percent Similarity: 90.28% Conservative: 4
Best Local Similarity: 88.43% Mismatches: 21
Query Match: 73.54% Indels: 0

DB: 3 Gaps: 0

US-09-441-857-12 (1-254) x AAZ93837 (1-740)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
 Db 71 ATGTCCACGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG 130

Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40
 Db 131 GAGCAAAAGCGTTGGAAAGACATCTTTGATCACCAGATTCAATGATGACAGTTTTCGACAC 190

Qy 41 ThrTyrGlnAlaIleIleGlyLysPheLeuSerLysThrMetTyrIleuGluAspGly 60
 Db 191 ACCTATCAGGCAACAATTTGGCATTCGACCTTTTATCAAAAACCTATGTTCTGGAGGATCGA 250

Qy 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
 Db 251 ACAGTACGATTGCAATTTATGGACACAGAGTCAAGAGCGGTTCCAGGAGCTTGATTCCT 310

Qy 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
 Db 311 AGCTACATTCTGTGACTCCACTGTGSCAGTTGTTGTTATGATATCACAAATGTTAACTCA 370

Qy 101 PheGlnGlnThrThrLysTyrIleAspAspValArgThrGluArgGlySerAspValIle 120
 Db 371 TTCCAGCAAACTACAAAGTGGATTGATGATGTGCAGAACAGAGGAAGTGTATGTTATC 430

Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
 Db 431 ATCATGCTAGTAGGAAATAAACAGATCTTGCTGACAGAGCAAGTGTCAATTGAGGAG 490

Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
 Db 491 GGACAGAGGAAGCCAAAGAGCTGAATGTTATGTTTATGAAACTAGTGCAAAAGCTGGA 550

Qy 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
 Db 551 TACAATGTAAAGCAGCTCTTTCGAGCTGTAGCAGCAGCTTTGCCGGGAATGGAAAGCACA 610

Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
 Db 611 CAGCAGAGAGCAGAGAGATATGATTGACATAAAACTGGAAAAAGCCTCAGGAGCAACCA 670

Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThr 216
 Db 671 GTCAGTGNAGAGGCGTGTCTCTGCTAATGTCCTTAGTCACTCTTCAACC 718

RESULT 13
 AAZ93841
 ID ID AAZ93841 standard; DNA; 624 BP.
 XX
 AC AAZ93841;
 XX
 XX
 DT 29-AUG-2000 (first entry)
 XX
 DE RAB6C coding sequence.
 XX
 KW WTH3; RAB6; drug resistance; locus; loci; methylation; tumour; probe;
 KW antibody; immunogen; mutation; detection; therapy; human; ss.
 XX
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT CDS 1..624
 FT /*tag= a
 FT /product= "RAB6C"
 XX
 XX
 PN WO200029625-A1.
 PD
 XX
 PD 25-MAY-2000.
 PF
 PF 18-NOV-1999; 99WO-US027630.

[illegible]

Db 361 ATCATGCTAGTAGGAAATAAACACAGATCTTGCTGACAGAGCAAGTGTCAATTGAGGAG 420
 Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
 Db 421 GGAGAGAGGAAAGCCAAAGAGCTGAATGTTATGTTTATTGAACTAGTGTCAAAAGCTGGA 480
 Qy 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
 Db 481 TACAATGTAAAGCAGCTCTTCGAGCTGTAGCAGCAGCTTTGCCGGGAATGGAAGACACA 540
 Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThr 200
 Db 541 CAGGACAGAAGCAGAGAAGATGATTGACATAAACTGGAAGAAAGCCTCAGGAGCAACA 600
 Qy 201 ValSerGluGlyGlyCysSerCys 208
 Db 601 GTCAAGTGAAGGAGGCTGTTCTCTGC 624

RESULT 14

AAK52711

ID AAK52711 standard; cDNA; 1219 BP.

XX AC

XX AAK52711;

XX 06-NOV-2001 (first entry)

XX DT

XX DE Human polynucleotide SEQ ID NO 2240.

XX KW

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation; ss.

XX OS

XX Homo sapiens.

XX PN

XX WO200157190-A2.

XX XX

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US004098.

XX PR 03-FEB-2000; 2000US-00496914.

XX PR 27-APR-2000; 2000US-00560875.

XX PR 20-JUN-2000; 2000US-00598075.

XX PR 19-JUL-2000; 2000US-00620325.

XX PR 01-SEP-2000; 2000US-00654936.

XX PR 15-SEP-2000; 2000US-00663561.

XX PR 20-OCT-2000; 2000US-00693325.

XX PR 30-NOV-2000; 2000US-00728422.

XX XX

XX PA (HYSEQ-) HYSEQ INC.

XX XX

XX PI Tang YT, Liu Q, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX XX

XX DR WPI; 2001-476283/51.

XX DR P-PSDB; AAM79578.

XX XX

XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX PT in diagnosis and gene therapy.

XX PS

PS Claim 1; Page 4593; 6221pp; English.

XX XX

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication

XX SQ Sequence 1219 BP; 287 A; 314 C; 337 G; 281 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,48e-92 Length: 1219
 Score: 892.00 Matches: 172
 Percent Similarity: 90.38% Conservative: 16
 Best Local Similarity: 82.69% Mismatches: 20
 Query Match: 68.40% Indels: 0
 DB: 4 Gaps: 0

US-09-441-857-12 (1-254) x AAK52711 (1-1219)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
 Db 179 ATGTCGCGAGGGGAGATTTTGGGAATCCACTGAGAAAATTCAGTTGGTGTCTTGGGG 238
 Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40
 Db 239 GAGCAGAGCGTCGGGAAGACGCTCTCTGATTACGAGGTTTCATGTACGACAGCTTCGACAAC 298
 Qy 41 ThrTyrGlnAlaIleIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
 Db 299 ACATACCAGGCAACCATTTGGGATTGACTTCTTGTCAAAAACCATGTACTTTGGAGGCGC 358
 Qy 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
 Db 359 ACGTGGCAGCTGCAGCTCTGGGACACAGCTGTCAGGAGAGGTTCCGCGAGCTGATCCCC 418
 Qy 81 ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAsnSer 100
 Db 419 AGCTACATCCGGGACTCCACGGTGGCTGTGGTGTACGACATCACAAATCTCAACTCC 478
 Qy 101 PheGlnGlnThrThrLysTyrIleAspAspValArgThrGluArgGlySerAspValIle 120
 Db 479 TTCCACACAGACCTCTAAGTGGATCGACGCTCAGGACAGAGAGGGGCGAGTGTATTTC 538
 Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
 Db 539 ATCATGCTGGTGGCAACAGACGACCTGCTGTATAGAGGACAGATACCATCGAGAG 598
 Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
 Db 599 GGGGAGCAGCGCGCCAAAGAACTGAGCGTCTATGTTTCATTGAGACCAGTGCAGAGACTGGC 658
 Qy 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
 Db 659 TACAACGTGAAGCAGCTTTTTCGACGCTGGCGCTCTACCCGGAATGGAGAATGTC 718
 Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThr 200
 Db 719 CAGGAGAAAGCAAGAGGATGATTGACATCAAGCTGGACAAACCCAGGAGCCCCCG 778
 Qy 201 ValSerGluGlyGlyCysSerCys 208
 Db 779 GCCAGCGAGGGCGCTGCTCTCTGC 802

RESULT 15

ADP07320

ID ADP07320 standard; DNA; 1266 BP.

XX AC

XX ADP07320;

XX XX

XX DT 29-JUL-2004 (first entry)

XX DE Human RAB6B DNA.

XX XX

XX KW ds; proliferative disease; breast; methylation; CpG; bisulfite; human.

OS Homo sapiens.
XX DE10255104-A1.
XX 11-MAR-2004.
XX
XX 26-NOV-2002; 2002DE-01055104.
XX 27-AUG-2002; 2002DE-01039311.
XX (EPITG-) EPIGENOMICS AG.
XX Maier S;
XX WPI: 2004-284340/27.
XX GENBANK; NM_016577.
XX
XX Analyzing proliferative diseases of breast cells, useful e.g. for
XX diagnosis, prognosis and treatment of breast cancer, by determining
XX methylation status of specific genes.
XX
XX Claim 1; Page; 22pp; German.
XX
XX This invention describes a novel method of analysing proliferative
XX diseases of breast cells by determining the methylation status of certain
XX genes. The invention also describes nucleic acids or their complements,
XX oligomers, especially oligonucleotides or peptide nucleic acid oligomers,
XX that hybridise to, or are identical with, any of the nucleic acids, the
XX preparation of an oligomer array for analysing proliferative diseases of
XX breast cells that are associated with the methylation status of CpG
XX dinucleotides of the genes by bonding at least one oligomer to a solid
XX phase and a kit comprising a bisulfite reagent and the oligomer. The
XX nucleic acids are genomic sequences (5'- and/or regulatory and/or CpG-
XX rich regions). The base sequence of the oligomer includes at least one
XX CpG island, especially with C in the middle third of the sequence. The
XX process involves a genomic DNA sample treated chemically, specifically
XX with a bisulfite reagent, to convert non-methylated C to uracil or some
XX other base having base-pairing properties different from those of C.
XX Fragments of the treated DNA are then amplified, using primers and a
XX polymerase and the methylation status of the genomic CpG dinucleotides is
XX determined by analysis of the amplicons, particularly by hybridisation to
XX the oligomer, optionally with extension of the hybridised oligomer by at
XX least one base, or detection is by sequencing. The amplification may use
XX methylation-specific primers. Alternatively, genomic DNA is extracted
XX from a sample and digested with methylation-specific restriction enzymes,
XX then the digestion fragments detected, optionally after amplification. In
XX either method, more than 10 fragments of 100-200 bp are amplified in a
XX single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
XX amplicons carry detectable markers, e.g. fluorophores, radioisotopes
XX and/or releasable fragments of known mass that can be detected by mass
XX spectrometry. The method is used for characterisation, classification,
XX differentiation, staging, phase-estimation, diagnosis and/or therapy of
XX proliferative diseases of breast cells. The method provides very specific
XX classification of proliferative diseases, allowing better treatment. It
XX can both characterise methylation status and detect single-nucleotide
XX polymorphisms. This sequence represents human gene used to illustrate the
XX method of the invention. NOTE: This sequence does not appear in the
XX printed specification but has been retrieved from Genbank.
XX
SQ Sequence 1266 BP; 269 A; 379 C; 344 G; 274 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6,84e-92 Length: 1266
Score: 892.00 Matches: 172
Percent Similarity: 90.38% Conservatives: 16
Best Local Similarity: 82.69% Mismatches: 20
Query Match: 68.40% Indels: 0
DB: 12 Gaps: 0

US-09-441-857-12 (1-254) x ADP07320 (1-1266)

QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLeuValPheLeuGly 20
|||||

Db 338 ATGTCGCGAGGGGGAGATTTTGGGAATCCACTGAGAAAATTCAGTTGGTGTCTTGGGG 397
QY 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAen 40
|||
Db 398 GAGCAGACGTCGGGAGACGCTCTCTGATTACGAGGTTTCATGTACGACAGCTTCGACAAAC 457
QY 41 ThrTyrGlnAlaIleIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
|||
Db 458 ACATACCAGGCAACCCATTGGGATTGACTTCTTGTCAAAAACCATGTACTTGGAGGACCGC 517
QY 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
|||
Db 518 ACGGTGCGACTGCGACACTCTGGGACACAGCTGGTCAGGAGAGGTTCGCGAGCTGTATCCCC 577
QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
|||
Db 578 AGCTACATCCGGGACTCCACGGT 637
QY 101 PheGlnGlnThrThrLysThrLysThrLysThrLysThrLysThrLysThrLysThrLys 120
|||
Db 638 TTCCAACAGACCTCTTAAGTGGATCGACGACGTTCAGGACAGAGAGGGGCGAGTGTATC 697
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
|||
Db 698 ATCATGTCTGTGGGCAACAAGACGACCTGGCTGATAGAGGCGAGATAACCATCGAGGAG 757
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
|||
Db 758 GGGGAGCAGCGCGCCCAAGAACTGAGCGTCTATGTTGAGACCCAGTGGGAAGACTGGC 817
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLysLeuProGlyMetGluSerThr 180
|||
Db 818 TACAACGTGAAGCAGCTTTTCGACGTGTGGCGTCTTACCCGGAATGGAGAATGTC 877
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
|||
Db 878 CAGGAGAAAGCAAGAGAGGATGATCATCAAGCTGGACATCAAGCTGGACAAACCCGAGGAGCCCG 937
QY 201 ValSerGluGlyGlyCysSerCys 208
Db 938 GCCAGCAGGCGGCTGTCTCTGTC 961

Search completed: April 25, 2005, 03:56:04
Job time : 609 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2005, 03:04:50 ; Search time 170 Seconds
(without alignments)
2444.790 Million cell updates/sec

Title: US-09-441-857-12
Perfect score: 1304
Sequence: 1 MSAGDGFNGPLKFKLVFLG.....NLFPSLITFCNSLLPVSWR 254

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USFTO_spool_p/US09441857/runat 22042005 155509 24777/app query.fasta_1.391
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09441857@cgn1 1.105 @runat 22042005 155509 24777 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfile1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1138	87.3	3745	4	Sequence 387, App
2	973	74.6	739	4	Sequence 2946, App
3	479	36.7	329	4	Sequence 6279, App
4	401.5	30.8	414	4	Sequence 6213, App
5	396	30.4	1546	4	Sequence 5540, App
6	391.5	30.0	719	4	Sequence 3023, App
7	388.5	29.8	1984	4	Sequence 7, Appli
8	377	28.9	2874	4	Sequence 2160, App
9	371.5	28.5	1630	4	Sequence 409, App
10	354.5	27.2	921	4	Sequence 1124, App
11	347.5	26.6	773	4	Sequence 4216, App
12	346.5	26.6	717	4	Sequence 5859, App

13	344	26.4	834	4	US-09-949-016-3883	Sequence 3883, Ap
14	342.5	26.3	1631	4	US-09-620-312D-587	Sequence 587, App
c 15	338	25.9	998	4	US-09-270-767-1882	Sequence 1882, Ap
c 16	338	25.9	998	4	US-09-270-767-17164	Sequence 17164, A
17	337.5	25.9	714	4	US-09-248-796A-5886	Sequence 5886, Ap
18	337	25.8	723	4	US-09-016-434-1422	Sequence 1422, Ap
19	335.5	25.7	1775	4	US-09-949-016-4926	Sequence 4926, App
20	335	25.7	1069	4	US-09-620-312D-646	Sequence 646, App
21	333	25.5	1255	4	US-09-949-016-1772	Sequence 1772, Ap
22	333	25.5	8137	4	US-09-566-921-7	Sequence 7, Appli
23	331.5	25.4	848	3	US-08-741-411-2	Sequence 2, Appli
24	331.5	25.4	920	4	US-09-949-016-4287	Sequence 4287, Ap
25	328.5	25.2	833	4	US-09-620-312D-426	Sequence 426, App
26	325.5	25.0	1053	4	US-09-270-767-12052	Sequence 12052, A
27	322.5	24.7	875	3	US-09-075-454-10	Sequence 10, Appl
28	322.5	24.7	1106	4	US-09-620-312D-959	Sequence 959, App
29	322.5	24.7	2612	3	US-09-484-970B-142	Sequence 142, App
30	322	24.7	639	3	US-09-399-913-66	Sequence 66, Appl
31	322	24.7	639	4	US-09-350-614-66	Sequence 66, Appl
32	321.5	24.7	809	4	US-09-949-016-3376	Sequence 3376, Ap
33	321.5	24.7	1148	4	US-09-949-016-4879	Sequence 4879, Ap
c 34	321.5	24.7	24257	4	US-09-949-016-13902	Sequence 13902, A
35	318.5	24.4	925	2	US-08-916-901-4	Sequence 4, Appli
36	318.5	24.4	925	3	US-09-154-602-4	Sequence 4, Appli
37	316.5	24.3	2552	4	US-09-270-767-15247	Sequence 15247, A
38	314.5	24.1	842	4	US-09-255-920A-6	Sequence 6, Appli
39	309.5	23.7	3936	4	US-09-919-172-49	Sequence 49, Appl
40	306.5	23.5	970	3	US-08-888-077A-28	Sequence 28, Appl
41	306.5	23.5	2334	4	US-09-949-016-1758	Sequence 1758, Ap
42	306	23.5	1275	4	US-09-949-016-3508	Sequence 3508, Ap
43	306	23.5	1799	4	US-09-774-528-36	Sequence 36, Appl
44	304.5	23.4	684	4	US-09-248-796A-6188	Sequence 6188, Ap
45	303	23.2	1511	4	US-09-949-016-1134	Sequence 1134, Ap

ALIGNMENTS

RESULT 1
US-09-976-594-387
; Sequence 387, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 387
; LENGTH: 3745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 107569.15
US-09-976-594-387

Alignment Scores:			
Pred. No.:	1.9e-134	Length:	3745
Score:	1138.00	Matches:	225
Percent Similarity:	89.32%	Conservative:	7
Best Local Similarity:	87.21%	Mismatches:	22
Query Match:	87.27%	Indels:	4
DB:	4	Gaps:	1

US-09-441-857-12 (1-254) x US-09-976-594-387 (1-3745)

Oy 1 MetSerAlaGlyGlyAspPheClyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
|||||

```
Db 494 ATGTCCAGCGGGAGAGCTTCGGGAATCCGCTGAGGAAATTCAGAGCTGGTGTCTCTGGGG 553
Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40
Db 554 GAGCAAGCGTGGAAAGACATCTTTGATCACCAGATTTCATGTATGACAGATTTTGACAAC 613
Qy 41 ThrTyrGlnAlaIleIleGlyLeuSerPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 614 ACCTATCAGGCACAACTTGGCATTGACCTTTTATCAAAACATATGATCTTGAGGATCGA 673
Qy 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 674 ACAGTACGATTGCAATTTATGGGACACAGCAGGTCAAGAGCGTTTCAGGAGCTTGATTCCT 733
Qy 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
Db 734 AGCTACATTCGTGACTCCACTGTGCAGTGTGTGTTTATGATATCACAAATGTTAACTCA 793
Qy 101 PheGlnGlnThrThrLysThrIleAspAspValAlaArgThrGluArgGlySerAspValIle 120
Db 794 TTCAGCAAACTACAAAGTGGATTGATGTCAGACAGAAAGAGGAAGTGTATGTTATC 853
Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 854 ATCATGCTAGTAGGAATAAAACAGATCTTGCTGACAGAGGCAAGTGTCAATTGAGGAG 913
Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 914 GGAGAGAGAAAGCCAAAGAGCTGAATGTTATGTTTATGAAACTAGTGCAGAAAGCTGGA 973
Qy 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 974 TACATGTAAAGCAGCTCTTCGAGGTGAGCAGAGCTTTGCCGGGAATGGAAGACACA 1033
Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 1034 CAGGACAGAGCAGAGAGATATGATTCACATAAACTGGAAAGCCTCAGAGCAACCA 1093
Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerThrLeuProGlnLys 220
Db 1094 GTCAGTGAAGGAGGCTGTCTCTGCTTAATCTCCCATGTCATCTTCAACCTCTTCAGAAC 1153
Qy 221 -----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
Db 1154 TCACGTCTTTGGCCCCCTTACTCTTTCAATGACGTGAGTGAATATGGCTTGAACCTT 1213
Qy 237 PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 1214 TTCCTTCAGTAATAACGTAATGCAATTCATCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 1267

RESULT 2
US-09-949-016-2946
; Sequence 2946, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2946
; LENGTH: 739
; TYPE: DNA
; ORGANISM: Human
```

```
US-09-949-016-2946
Alignment Scores:
Pred. No.: 1,7e-114 Length: 739
Score: 973.00 Matches: 193
Percent Similarity: 91.20% Conservative: 4
Best Local Similarity: 89.35% Mismatches: 19
Query Match: 74.62% Indels: 0
DB: 4 Gaps: 0

US-09-441-857-12 (1-254) x US-09-949-016-2946 (1-739)
Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 70 ATGTCACGGCGGAGACTTCGGGAATCCGCTGAGGAAATTCAGAGCTGGTGTCTCTGGGG 129
Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40
Db 130 GAGCAAGCGTGGAAAGACATCTTTGATCACCAGATTTCATGATGACAGATTTTGACAAC 189
Qy 41 ThrTyrGlnAlaIleIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 190 ACCTATCAGGCACAAATTTGCAATTTTATCAAAAACTATGTACTTTGGAGGATCGA 249
Qy 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 250 ACAGTACGATTGCAATTTATGGACACAGCAGGTTCAGAGCGTTTCAGAGCTTGATTCCT 309
Qy 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
Db 310 AGCTACATTCGTGACTCCACTCGGCGAGTTGTTGTTATGATATCACAAATGTTAACTCA 369
Qy 101 PheGlnGlnThrThrLysThrIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 370 TTCCAGCAAACTACAAAGTGGATTGATGATGTCAGAACAGAGGCAAGTGTCAATTTGAGGAG 429
Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 430 ATCATGCTAGTAGGAATAAAACAGATCTTGCTGACAGAGGCAAGTGTCAATTTGAGGAG 489
Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 490 GGAGAGAGGAAAGCCAAAGAGCTGAATGTTATGTTTATTGAAACTAGTGCAGAAAGCTGGA 549
Qy 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 550 TACAAATGAAAGCAGCTCTTTTCGACGTGAGCAGCAGCTTTGCCGGGAATGGAAGACACA 609
Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 610 CAGGACAGAGCAGAGAGATATGATTGACATAAACTGGAAAGCCTTCAGGAGCAACCA 669
Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerThr 216
Db 670 GTCAGTAGAGGAGCTGTCTCTGCTTAATCTCCCATGTCATCTTCAACC 717

RESULT 3
US-09-513-999C-2679
; Sequence 2679, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
```

```

US-09-441-857-12 (1-254) x US-09-248-796A-6213 (1-414)
Qy 6 AspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGlyGlnSerValAla 25
Db 79 GATAAATCAATTTACTTAAATAAACAAGATGCTCTTCTTGGTGATCAAAAGTTGGT 138
Qy 26 LysThrSerLeuIleThrArgPheArgTyraAspSerPheAspAsnThrTyrGlnAla 45
Db 139 AAAACATCATTAATCACCAGATTTATGTATGATACATTTGATGAACACTTATGCTGCCACG 198
Qy 46 IleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly---ThrIleGlyLeu 64
Db 199 ATTGGAATGATTTTATCGAATAAATGATTTAGAGAAGGTAAACCATAGATTATTA 258
Qy 65 ArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArg 84
Db 259 CAATTATGGGATCTGCGGACACAGAAAGATTTTCGATCATTAATACCTTCATATATTAGA 318
Qy 85 AspSerAlaAlaValValTyrAspIleThrAsnValAsnSerPheGlnGlnThr 104
Db 319 GATTCATGTTGTCAGTAAATGTTATGATATAACCAATAAAAAATCATTTTGATAATCTT 378
Qy 105 ThrLysTrpIleAspAspValArgThrGluArgGly 116
Db 379 GATAAATGCTAAAGATGTTAAATTAGAACGAGGT 414

RESULT 5
US-09-949-016-5540
; Sequence 5540, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5540
; LENGTH: 1546
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5540

Alignment Scores:
Pred. No.: 3,228-40 Length: 1546
Score: 396.00 Matches: 83
Percent Similarity: 57.80% Conservative: 43
Best Local Similarity: 38.07% Mismatches: 86
Query Match: 30.37% Indels: 6
DB: 4 Gaps: 2

US-09-441-857-12 (1-254) x US-09-949-016-5540 (1-1546)
Qy 8 GlyAsnProLeuArgLysPheLysLeuValPheLeuGlyGlnSerValAlaYthr 27
Db 134 GGAACAAGATCTGCAATTTAAGCTGTTCTGCTGGGGAGATCTGCGTAGGCAATTC 193
Qy 28 SerLeuIleThrArgPheArgTyraAspSerPheAsnThrTyrGlnAlaIleGly 47
Db 194 AGCTCGTCTCCGCTTGTCAAGGACAGTTCACGAGTACCAGGAGACCAATTGGA 253
Qy 48 IleAspPheLeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTrp 67
Db 254 GGGGCGCTTCTCACACAGACTGCTCTGCTGGATGATGACACACAGTCTTGGATCTGG 313

US-09-441-857-12 (1-254) x US-09-513-999C-2679 (1-329)
Qy 5 GlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGlyGlnSerVal 24
Db 12 GGAGATTTTGGGAATCCACTGAGAAATTCAGATTGGTCTTGGGGGACAGCGGTC 71
Qy 25 AlaLysThrSerLeuIleThrArgPheArgTyraAspSerPheAspAsnThrTyrGlnAla 44
Db 72 GGGAGACGCTCTGATTACGAGTTTCATGACGACAGCTTCGACCAACATACACGCA 131
Qy 45 IleIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeu 64
Db 132 ACCATTGGGATTCATCTTGTCAAAACCATGTTACTTGGAGGACCGCACGGTGGCGACTG 191
Qy 65 ArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArg 84
Db 192 CAGCTCTGGGACACAGCTGCTGAGAGAGGTTCCGACGCTGATCCCGAGTACATCCGG 251
Qy 85 AspSerAlaAlaValValTyrAspIleThrAsnValAsnSerPheGlnGlnThr 104
Db 252 GACTCCACGGTGGCTGTGGTGTGTACGACATCACAAATCTCACTCTTCCACAGACC 311
Qy 105 ThrLysTrpIleAsp 109
Db 312 TCTAAGTGGATCGAC 326

RESULT 4
US-09-248-796A-6213
; Sequence 6213, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6213
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6213

Alignment Scores:
Pred. No.: 7,578-42 Length: 414
Score: 401.50 Matches: 80
Percent Similarity: 79.46% Conservative: 9
Best Local Similarity: 71.43% Mismatches: 22
Query Match: 30.79% Indels: 1
DB: 4 Gaps: 1

```


TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1984 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: U937NOT01
CLONE: 000513
US-09-023-655-7

Alignment Scores:

Pred. No.: 4,38e-39 Length: 1984
Score: 388.50 Matches: 82
Percent Similarity: 58.8% Conservative: 41
Best Local Similarity: 39.23% Mismatches: 81
Query Match: 29.79% Indels: 5
DB: 4 Gaps: 2

US-09-441-857-12 (1-254) x US-09-023-655-7 (1-1984)

Qy 8 GlyAsnProLeuArgGlyPheLysLeuValPheLeuGlyGluGlnSerValAlaLysThr 27
Db 11 GGAAATAAAATATGCCAGTTCAACTAGTACTTCTGGGAGAGTCCGCTGTTGGCAATCA 70
Qy 28 SerLeuIleThrArgPheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGly 47
Db 71 AGCCTAGTGCCTTCGTTTGTGAAGGCCCAATTCATGAATTTCAAGAGAGTACCATTGGG 130
Qy 48 IleAspPheLeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTrp 67
Db 131 GCTGCTTTTCAACCAACTGTATGCTCTGTATGACACTACAGTAAAGTTTCAATATGG 190
Qy 68 AspThrAlaGlyGlnGluArgSerLeuIleProArgTyrIleArgAspSerAla 87
Db 191 GATACAGCTGGTCAAGACGATACCATAGCTAGCACCCTAGTACCAATGTACTACAGAGGACACAA 250
Qy 88 AlaAlaValValTyrAspIleThrAsnValAsnSerPheGlnGlnThrThrLysTrp 107
Db 251 GCAGCCATAGTTGTATATGATATCAATAATGAGGAGTCTTTTGCAGAGACCAAAAAATTGG 310
Qy 108 IleAspAspValArgThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArg 127
Db 311 GTTAAAGAACTTCAGAGGCAACCAAGTCTCACTATGTAATGATCTTTATCGGAAACAAAG 370
Qy 128 ThrAspLeuAlaAspLysArgGlnValSerValGluGlyGluArgLysAlaLysGly 147
Db 371 GCCGACCTAGCAAAATAAAGACGAGTATGATTTCCAGGAAGCACAGTCTATGCAGATGAC 430
Qy 148 LeuAsnValThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPhe 167
Db 431 AATAGTTTATTATTCATGGAGACATCCGCTAAACATCAATGAATGTAATGAAATATTC 490
Qy 168 ArgArgValAlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAsp 187
Db 491 ATGGCAATAGCTAAAAAATTGCCAAGAAATGAACCAAAATCCAGGAGCAAAATTCCTGCC 550
Qy 188 MetSer---AspIleLysLeuGlyProGlnGlnThrValSerGluGlyCys 206
Db 551 AGAGGAAGAGGAGTAGACCTTACCGAACCACACAAACCAACCAAGGAATCAG----- 601
Qy 207 SerCysTyrSerProMetSerSerSer 215
Db 602 ---TGTGTGTAGTAACCTTCTAGT 625

RESULT 8

US-09-949-016-2160
Sequence 2160, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIORITY FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2160
LENGTH: 2874
TYPE: DNA
ORGANISM: Human
US-09-949-016-2160

Alignment Scores:

Pred. No.: 2,35e-37 Length: 2874
Score: 377.00 Matches: 91
Percent Similarity: 49.44% Conservative: 41
Best Local Similarity: 34.08% Mismatches: 105
Query Match: 28.91% Indels: 30
DB: 4 Gaps: 4

US-09-441-857-12 (1-254) x US-09-949-016-2160 (1-2874)

Qy 13 LysPheLysLeuValPheLeuGlyGluGlnSerValAlaLysThrSerLeuIleThrArg 32
Db 141 CAGTTCAAATTTGGTCTCTGGGAGATCTCTCAGTGGGAAAGTCAAGCTGCTGTATTACGT 200
Qy 33 PheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGlyLysPheLeuSer 52
Db 201 TTTGTCAAGGGCGAGTCTCCATGATGACAGGAGACACCATTTGGAGCGGCTTCTCTACC 260
Qy 53 LysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTrpAspThrAlaGlyGln 72
Db 261 CAGTCCGTTTCTTAGATGACACACAGTGAAGTTTGAGATCTGGGACACAGCTGGGCGAG 320
Qy 73 GluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaValValVal 92
Db 321 GAGCGATATCACAGCTTAGCCCTTACATGATCTACAGGGGTGCGCAAGCTGCAATCGTGT 380
Qy 93 TyrAspIleThrAsnValAsnSerPheGlnGlnThrThrLysTrpIleAspAspValArg 112
Db 381 TACGACATTACTAATCAGAAACCTTTGCCGAGCAAGACATGGGTGAAGAACTACAG 440
Qy 113 ThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeuAlaAsp 132
Db 441 CGACAGGCGCATCTAGCATCGTTATTCCTCTGGCAGGGAACAAAGCTGACCTGGCCCAAC 500
Qy 133 LysArgGlnValSerValGluGlyGluArgLysAlaLysGlyLeuAsnValThrPhe 152
Db 501 AAACGTATGTGTGATATGAAGAGCCCGCATATGCAGATGCACACAGCTTATTGTTC 560
Qy 153 IleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgValAlaAla 172
Db 561 ATGGAGACTTCAGCCCAAGACAGCTATGAACGTGAATGATCTCTCTCGCAATAGCTAAG 620
Qy 173 AlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAspMetSer---AspIle 191
Db 621 AAGTTGCCAAAGATGAACCCAGATCTGGGAGTGCAGAGCGGCGGAGCGGGGTGTG 680
Qy 192 LysLeuGluLysProGlnGlnThrValSerGlu-----GlyGly 205
Db 681 GATCTCCATGAACAGTCCAGCAGCAAGAGCCAGTGTGTAGCAACTAGAGGGGTGCG 740
Qy 206 -----Cys 206
Db 741 TAGCAGCAAAACAGTATGAGCTAGCACAGAGCTAAGAAATAACCTCCATCCCTACCCC 800

```

QY 207 SerCysTyrSerProMetSerSerThrLeuProGln-----LysProPro 222
Db 801 TCAGCACAACCCCTACGGTAACAGCACATGACCTGGCTCCCAAGGCGTGCCTCCT 860
QY 223 TyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeuIleThr 242
Db 861 GACAGCTCGGTCATGCGACTTTTAAACGCTTCAGCAACAACACAGGCGAGCTGTGCGCA 920
QY 243 PheCysAsnSerSerLeuLeu 249
Db 921 CTGGCTCTCTACCCCTACTC 941

RESULT 9
US-09-949-016-409
; Sequence 409, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 409
; LENGTH: 1630
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-409

Alignment Scores:
Pred. No.: 4,75e-37 Length: 1630
Score: 371.50 Matches: 77
Percent Similarity: 58.33% Conservative: 35
Best Local Similarity: 40.10% Mismatches: 79
Query Match: 28.49% Indels: 1
DB: 4 Gaps: 1

US-09-441-857-12 (1-254) x US-09-949-016-409 (1-1630)

QY 13 LysPheLysLeuValPheLeuGlyGluGlnSerValAlaLysThrSerLeuIleThrArg 32
Db 78 CAGTTCAATTTGGTCTCTGGGAGATCTCGAGTGGGAAAGTCAAGCTGTGTATTACGT 137
QY 33 PheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGlyIleAspPheLeuSer 52
Db 138 TTTTGAAGGCGAGTTCATCAGTACAGGAGACCATTTGGAGCGCTTCTCTACC 197
QY 53 LysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAlaGlyGln 72
Db 198 CAGTCCGTTTCTAGATGACACACAGTGAAGTTTGAGATCTGGGACACAGCTGGGCGAG 257
QY 73 GluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaValValVal 92
Db 258 GAGCGATATCACAGTTAGCCCTTACGAGGTGTACTACAGGGGTGCCCAAGCTGCAATCGTGT 317
QY 93 TyrAspIleThrAsnValAsnSerPheGlnGlnThrTyrLysTrpIleAspAspValArg 112
Db 318 TACGACATTACTAATCAGAAACCTTTGCCGAGCAAGACATGGTGAAGGAAGTACAG 377
QY 113 ThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeuAlaAsp 132
Db 378 CGACAGGCGAGCTCCTAGCATCGTTATTTCCTGGCAGGGAACAAAGCTGACCTGGCCAAC 437
QY 133 LysArgGlnValSerValGluGluGlyAlaLysAlaLysGlyLeuAsnValThrPhe 152

```

```

Db 438 AAACGTATGGTGGAGTATGAGAGCGCCAGCATATGCAGATGACACAGCTTATTGTTTC 497
QY 153 IleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgValAlaAla 172
Db 498 ATGAGAGACTTCAGCCAAAGACAGCTATGAACGTGAATGATCTCTCTCTGGCAATAGCTAAG 557
QY 173 AlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAspMetSer--AspIle 191
Db 558 AAGTTGCCAAAGAGTGAACCCAGAAATCTGGAGGTGCAGCGAGCCGAAAGCCGGGTGTG 617
QY 192 LysLeuGluLysProGlnGluGlnThrValSerGlu 203
Db 618 GATCTCCATGAACAGTCCAGCAGAACAGAGCCAG 653

RESULT 10
US-09-016-434-1124
; Sequence 1124, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA: US/09/016,434
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1124:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: GI388194
US-09-016-434-1124

Alignment Scores:
Pred. No.: 2,81e-35 Length: 921
Score: 354.50 Matches: 67
Percent Similarity: 60.32% Conservative: 47
Best Local Similarity: 35.45% Mismatches: 72
Query Match: 27.19% Indels: 3
DB: 4 Gaps: 1

US-09-441-857-12 (1-254) x US-09-016-434-1124 (1-921)

```


QY 11 LeuArgLysPheLysLeuValPheLeuGlyGluInSerValAlaLysThrSerLeu1le 30
Db 67 ATACGGAGCTCAAGATGTCCTTCGGGGACACTGGGGTTGGGAATCAAGCATCGTG 126
QY 31 ThrArgPheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleGlyLeuAspPhe 50
Db 127 TGTGATTTGTCAGGATGACATTTTACCAACAACATCAGCCCTTACTTGGGGCACTTTT 186
QY 51 LeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAla 70
Db 187 ATGACCAAAACGTGCTTGTGGAAATGAACCTTCAAGATTCCTCATCTGGGACACTGCT 246
QY 71 GlyGlnGluArgLeuSerLeu1leProArgTyrIleArgAspSerAlaAlaVal 90
Db 247 GGTCAAGAACGGTTTCATTCATTTGCTCCCATGTACTATCGAGGCTCAGCTCAGCTGT 306
QY 91 ValValTyrAspIleThrAsnValAsnSerPheGlnInThrThrLysTrpIleAspAsp 110
Db 307 ATCGGTATGATATTACCAAGCAGGATTCATTTTATACCTTGAAGAAATGGGTCAAGGAG 366
QY 111 ValArgThrGluArgGlySerAspValIleThrLeuValGlyAsnArgThrAspLeu 130
Db 367 CTGAAGAACATGCTCCAGAAACATTTGTAATGGCCATCGCTGGAAACAAGTGCACCTC 426
QY 131 AlaAspLysArgGlnValSerValGluGluGlyGluArgLysAlaLysGlyLeuAsnVal 150
Db 427 TCAGATATTAGGAGGTTCCCTGAAGGATGCTAAGGAATACGCTGAATCCATAGTGC 486
QY 151 ThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgVal 170
Db 487 ATCGGGTGTGACAAAGTGCACAAATATGCTATTATATATCGAAGACTCTTTCAAGGAATC 546
QY 171 AlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArgLysAspMetSerAsp 190
Db 547 AGCCCCAGATCCACCTTGACCCCATGAAATGAAACAAT-----GGAACA 597
QY 191 IleLysLeuGluLysProGlnGluGln 199
Db 598 ATCAAAGTTGAGAGCAACCAACCATGCAA 624

RESULT 11

US-09-949-016-4216
; Sequence 4216, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4216
; LENGTH: 773
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4216

Alignment Scores:
Pred. No.: 1.66e-34 Length: 773
Score: 347.50 Matches: 79
Percent Similarity: 50.67% Conservative: 35
Best Local Similarity: 25.11% Mismatches: 86
Query Match: 26.65% Indels: 25
DB: 4 Gaps: 4

US-09-441-857-12 (1-254) x US-09-949-016-4216 (1-773)
QY 14 PheLysLeuValPheLeuGlyGluInSerValAlaLysThrSerLeu1leThrArgPhe 33
Db 102 TTCAAGATTCATCATCGGCAACAGCAGCGTGGCAGACGTCCTTCTCTTCGCGTAT 161
QY 34 ArgTyrAspSerPheAspAsnThrTyrGlnAlaIleGlyLeuAspPheLeuSerLys 53
Db 162 GCTCAGCACTCGTTTCAAGCTTCTCGTTCAGCAGCGTGGGCATCGACTTCAAGGTCAAG 221
QY 54 ThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGlu 73
Db 222 ACCATCTATCCCAACGACAGAGATCAAGCTCGAGATCTCGGACACACGAGGCGCAAGAG 281
QY 74 ArgLeuArgSerLeu1leProArgTyrIleArgAspSerAlaAlaValValValTyr 93
Db 282 CGGTACCGGACCATCACCCGCACTATACCGGGCGCTATGGGCTTCATCTCATGTAT 341
QY 94 AspIleThrAsnValAsnSerPheGlnInThrThrLysTrpIleAspAspValArgThr 113
Db 342 GACATCAACCAACGAGGAATCTTCAATGCAGTCAGGACTGCTCCACCACCATCAAGACC 401
QY 114 GluArgGlySerAspValIleThrLeuValGlyAsnArgThrAspLeuAlaAspLys 133
Db 402 TACTCATGGGCAATGCCAGTGTCTGCTAGGAAACAAGTGTGACATGGAGGATGAG 461
QY 134 ArgGlnValSerValGluGlyGluArgLysAlaLysGlyLeuAsnValThrPheIle 153
Db 462 CGGTGTGTCTCATCAGAACGTGGCGGCTAGTGTGACCATCTGGGTTCGAGTCTTT 521
QY 154 GluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgValAla----- 171
Db 522 GAGCAAGCGCAAGGACCAACATTAACTGTCAGCAGACCTTTGAGCGCTGTGTGATGTC 581
QY 172 -----AlaAlaLeuProGlyMetGluSerThrGln 181
Db 582 ATCTGCGAGAAAGATGTCGAGTCGTGGACACGCGGACCTGCGGTGCACAGCGCCCAAG 641
QY 182 AspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThrVal 201
Db 642 CAGGCGCCACAG-----CTCAGTGACACGAGGTGCCACCGCACCAGGAC----- 686
QY 202 SerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThr-----Leu 217
Db 687 -----TGGCGCTGTGAGAGCACTCCCACTTCCCTTTCCTCTTCCCTGTCTT 734
QY 218 ProGlnLysProPro 222
Db 735 CCCACCTTCCCCCA 749

RESULT 12

US-09-248-796A-5859
; Sequence 5859, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5859
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-5859

Alignment Scores:

```
Pred. No.: 1 97e-34 Length: 717
Score: 346.50 Matches: 75
Percent Similarity: 57.62% Conservative: 46
Best Local Similarity: 35.71% Mismatches: 72
Query Match: 26.57% Indels: 17
DB: 4 Gaps: 4

US-09-441-857-12 (1-254) x US-09-248-796A-5859 (1-717)
Qy 13 LysPheLysLeuValPheLeuGlyGluGlnSerValAlaLysThrSerLeuIleThrArg 32
Db 85 CAATTCAAACTTGCTTTAGGAGAAAGTCAGTGGTAAATCATCATCGTACACCGT 144
Qy 33 PheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGlyIleAspPheLeuSer 52
Db 145 TTTGTCAAGAACACATTTGATATATAGAGAAATCGACTATAGTGCAGCATTTTAAACA 204
Qy 53 LysThrMetTyrLeuGluAspGly-----ThrIleGlyLeuArgLeuTyrAspThrAla 70
Db 205 CAACATATACTATACCGGAATCCGAGACGAGATTAATAATTGAAATATGGATCTGCA 264
Qy 71 GlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaVal 90
Db 265 GGACAGGAGCGTTACAAGTCATTGGCACCACATGATTATAGAAAGCCCAATGCTGCA 324
Qy 91 ValValTyrAspIleThrAsnValAsnSerPheGlnGlnThrThrLysTyrIleAspAsp 110
Db 325 TGTGTTTATGATATACTAGTGTAGTCAATTCAAAAGGCCCAAGATTGGATCAAGAA 384
Qy 111 ValArgThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeu 130
Db 385 TTGAAAAGCAGCTCCAGAGAGTATTGTTATGCTCTAGTGGGAACAAGTCTGATT 444
Qy 131 AlaAspLysArgGlnValSerValGluGlu-----GlyGluArgLysAla 145
Db 445 GACGATGAAGAGAGAGTCAAGTCAAGTGAAGTCAAGTATGTTCAAGACCAAAATAGC 504
Qy 146 LysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGln 165
Db 505 GATGCTGTTCTATAATTACTGCTGAATGTTCTGCTAAATCTGGAGATGGGGTTTGGAT 564
Qy 166 LeuPheArgArgValAlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArg 185
Db 565 GTATTTAACAAATAGGAAGACITTTACCT---GTGGACGAAGTTATAGCTAGTAGA 621
Qy 186 GluAspMet-----SerAspIleLysLeuGluLysPro 196
Db 622 TCAAGACAAAGTGTGTTGCTGGAGGAAGAACACGAGGAGTGTGTTTGAATAGACCT 681
Qy 197 GlnGluGlnThrValSerGluGlyGlyCys 206
Db 682 AGAGGTCAAGTAAATCAAAGCAACTCTTTC 711

RESULT 13
US-09-949-016-3883
; Sequence 3883, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3883

; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3883

Alignment Scores:
Pred. No.: 5.26e-34 Length: 834
Score: 344.00 Matches: 74
Percent Similarity: 50.49% Conservative: 29
Best Local Similarity: 36.27% Mismatches: 85
Query Match: 26.38% Indels: 16
DB: 4 Gaps: 2

US-09-441-857-12 (1-254) x US-09-949-016-3883 (1-834)
Qy 14 PheLysLeuValPheLeuGlyGluGlnSerValAlaLysThrSerLeuIleThrArgPhe 33
Db 241 TTCAAACTGTACTGATAGGCAACAGCAGTGTGGCAAGACTTCTTCTCTGTTCCGATAC 300
Qy 34 ArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGlyIleAspPheLeuSerLys 53
Db 301 GCGGACGACTCTTCACTCCCGCTTCTGCTCAGTACTGTGGGCATCGATTTCAAGGTCAAG 360
Qy 54 ThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGlu 73
Db 361 ACCGTCTACCCGCCATGACAAGAGGATCAAGCTGCAGATCTGGGACACAGCGGGCAGAG 420
Qy 74 ArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaValValValTyr 93
Db 421 CGCTACCGCACCATCACCGCCCTACTACCGGGAGGCCATGGGCTTCTGCTCATGTAT 480
Qy 94 AspIleThrAsnValAsnSerPheGlnGlnThrThrLysTyrIleAspAspValArgThr 113
Db 481 GACATCGCCAATCAGGAATCTTTTGGCGCTGTGCAGGACTGGGCCACGCAAAATCAAGACC 540
Qy 114 GluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeuAlaAspLys 133
Db 541 TACTCTGGGACACACCCAGGCTCTCTGTGGGGAACAAGTGTGACCTGGAGGACGAA 600
Qy 134 ArgGlnValSerValGluGluGlyGluArgLysAlaLysGlyLeuAsnValThrPheIle 153
Db 601 CGTGTGCTGCTGTGAGGATGCGCGAGGCTGCCGACGACCTTGTGTTTCGAGTCTTT 660
Qy 154 GluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgValAlaAla 173
Db 661 GAAAGCCAGTCCCAAGGAGAACATCAATGTGAAGCAGGTCTTCGAGCGCTGGTGATGTC 720
Qy 174 Leu-----ProGlyMetGluSerThrGlnAspGlySer 184
Db 721 ATCTGCGAGAGATGAACGAGTCCCTGGAACCCAGCTCCAGCTCAGCAGCAACGGGAAA 780
Qy 185 ArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThrValSerGluGly 204
Db 781 GGCCCGCGCTGGGGATGTCTCAGCCCCCAGGCC-----AGC 819
Qy 205 GlyCysSerCys 208
Db 820 AGCTGCAGCTGC 831

RESULT 14
US-09-620-312D-587
; Sequence 587, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
```

```
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yuning
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ POLYPEPTIDES
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ PRIORITY FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIORITY FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIORITY FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 587
/ LENGTH: 1631
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (66)..(650)
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)....(1631)
/ OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-587

Alignment Scores:
Pred. No.: 2,42e-33 Length: 1631
Score: 342.50 Matches: 80
Percent Similarity: 49.43% Conservative: 50
Best Local Similarity: 30.42% Mismatches: 108
Query Match: 26.27% Indels: 25
DB: 4 Gaps: 3

US-09-441-857-12 (1-254) x US-09-620-312D-587 (1-1631)

Qy 3 AlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGlyGluIn 22
Db 48 GCCCTCGCTCCCGGCCATGCGCTGAGGAGCTCAAGTGTCTGCTCGGGGATACA 107
Qy 23 SerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsnThrTyr 42
Db 108 GGTGTAGTAAATCGATGATTGTGTGGCGGTTTGTGGAGACAGATTTTGTATCCAAACATC 167
Qy 43 GlnAlaIleIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGlyThrIle 62
Db 168 AACCCAAACATAGGGGCATCTTTATGACCAAGCTGCTCCAGTACCAAAATGAGGTACAT 227
Qy 63 GlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLysLeuIleProArgTyr 82
Db 228 AAATCTCTAATCTGGGATACAGCTGGCAAGAACGATTTTGGTCTTACCAACATGATAC 287
Qy 83 IleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSerPheGln 102
Db 288 TATCGAGGTCGGCTGCGACTATAATCGTTTATGATATCACAAAAGAGAGACATTTTCA 347
Qy 103 GlnThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIleIleThr 122
Db 348 ACATTAAGAATTTGGTGAAGAGCTTTCAGCAGCATGGCCACCTAAATATTGTAGTGGC 407
Qy 123 LeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlyGlu 142
Db 408 ATTGAGGAAATAAATGATCTTATCATGATGTAAGAGAAATCATGGAGAGAGATCCAAAG 467
Qy 143 ArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGlyTyrAsn 162
Db 468 GACTACGCCGACTCTATTATTCGAATTTTGTAGAGACCGCGCAAAACCGCGATAAAC 527

163 VallysGlnLeuPheArgValAlaAlaLeuPro----- 175
176 -----Gly-MetGluSerThrGlnAspGlySerArgGluAspMetSer-AspIleLysL 193
588 TCTGCGCGTAAGGGCTTCAAACTCCAGAGACAGCCTTCAGAGCCAAAGCGAGCTGCTGC 647
193 euGluLysProGlnGluInThrValSerGlu----- 203
648 TGACCGAACCTCAGCCTCTCAGACTTGATGATGAAGTAGGTAGTCTCCTGAAAGTTAACAGG 707
204 -----GlyGlyCysSerCysTyrSerProMetSerSerThrLeuProGlnLysP 221
708 AGGGCTGGGGTCCCTGCCACCAGTTTTCACCTAGGCAGTCTTGTAGTCTTCCCGGCAAA 767
221 roProTyrSerPheIleAsp-----CysSerValAsnIleGlyLeuAsnLeuPheProS 239
768 AAGGATTTCAGAAATGACCACTTCTGTTCTCCAAAGACTGCAGCAATGATATTTCAGT 827
239 erLeu 240
828 CTGTG 832

RESULT 15
US-09-270-767-1882/c
; Sequence 1882, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1882
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1882

Alignment Scores:
Pred. No.: 4.12e-33 Length: 998
Score: 338.00 Matches: 73
Percent Similarity: 57.64% Conservative: 44
Best Local Similarity: 35.96% Mismatches: 82
Query Match: 25.92% Indels: 4
DB: 4 Gaps: 2

US-09-441-857-12 (1-254) x US-09-270-767-1882 (1-998)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 883 ATGAGCTCGACAGAACGAGGAGCGGCCCGCTTAAATTTCAAGCGCGTGTCTGCTGGCG 824
Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAen 40
Db 823 GAAGGTTGTGTGGGACAGACGTCGTGCTGCGCTACATGGAAGACCGGTTCATATGCC 764
Qy 41 ThrTyrGlnAlaIleIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 763 CAGCACCTAACGACCCCTGCAGCTTCTTCTGACGCCGCAAGATGTCCTCGAGGATGGG 704
Qy 61 ---ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIle 79
Db 703 AGAAGGCGCAGTGTGAATATTGGGACACGCGTGTGTCAGGAGCGGTTCCACGCCCTGGGA 644
Qy 80 ProArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAen 99
Db 643 CCCATCTACTATCGCGCTCCGAGCGCCCTGCTGCTATGATACATAACGACCGACGAC 584
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2005, 04:55:10 ; Search time 6740 Seconds
(without alignments)
228.790 Million cell updates/sec

Title: US-09-441-857-12

Perfect score: 1304

Sequence: 1 MSAGDGFNPKFKLVFLG.....NLFLSLITFCNSLLPVSWR 254

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09441857/runat_22042005_155511_24863/app_query.fasta_1.391
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MTNMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09441857_@CGN_1_1_723_@runat_22042005_155511_24863
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C					
1	1299	99.6	744802	17	US-10-292-798-1369
2	1133	86.9	3118	9	US-09-925-300-424
3	959	73.5	740	17	US-10-116-275-157
4	892	68.4	1266	17	US-10-172-118-1656
5	892	68.4	1266	17	US-10-342-887-1656
6	892	68.4	1266	19	US-10-948-755A-181
7	846.5	64.9	1157	13	US-10-108-605-44
8	835.5	64.1	2456	17	US-10-094-749-430
9	715	54.8	926	18	US-10-653-047-7008
10	710	54.4	1175	17	US-10-437-963-63907
11	709.5	54.4	1200	17	US-10-425-114-4106
12	709.5	54.4	1220	17	US-10-425-114-1068
13	709.5	54.4	1506	18	US-10-425-115-20462
14	707	54.2	1242	18	US-10-425-115-20463
15	705.5	54.1	1308	17	US-10-425-114-27873
16	704.5	54.0	1145	18	US-10-425-115-160784
17	699.5	53.6	1069	17	US-10-425-114-27069
18	697	53.5	1317	17	US-10-424-599-47043
19	688	52.8	1155	17	US-10-424-599-47044
20	667	51.2	523	18	US-10-021-323-14369
21	662.5	50.8	714	17	US-10-424-599-47503
22	653	50.1	1045	18	US-10-437-963-43192
23	649.5	49.8	1040	17	US-10-425-114-28361
24	649.5	49.8	1116	18	US-10-425-115-181339
25	642	49.2	443	9	US-09-960-352-7355
26	623	47.8	1296	18	US-10-739-930-4235
27	596	45.7	791	18	US-10-767-795-6480
28	593.5	45.5	3826	15	US-10-017-161-1727
29	593.5	45.5	3826	17	US-10-292-798-1383
30	582	44.6	648	17	US-10-369-493-25486
31	572	43.9	846	17	US-10-369-493-25812
32	569	43.6	424	10	US-09-918-995-5540
33	562	43.1	714	18	US-10-653-047-213
34	541.5	41.5	5158	17	US-10-767-795-6479
35	536.5	41.1	5758	17	US-10-275-933-11
36	532	40.8	1072	17	US-10-424-599-72715
37	485	37.2	570	18	US-10-437-963-43190
38	467	35.8	525	17	US-10-369-493-36580
39	461.5	35.4	963	17	US-10-424-599-47042
40	431	33.1	320	9	US-09-867-701-8545
41	431	33.1	1042	17	US-10-260-238-3339
42	396	30.4	819	17	US-10-262-511-221
43	396	30.4	960	17	US-10-262-511-217
44	396	30.4	1007	9	US-09-822-849A-563
45	396	30.4	1529	17	US-10-205-331-71

ALIGNMENTS

RESULT 1

US-10-292-798-1369/c
; Sequence 1369, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR FILING DATE: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1369

```
; LENGTH: 744802
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: (310089)..(310188)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (310089)..(310188)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (332935)..(332935)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (332992)..(332992)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (362002)..(362101)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (639781)..(639880)
; OTHER INFORMATION: a, t, c, g, unknown or other
; US-10-292-798-1369
Alignment Scores:
Pred No.: 1.78e-151 Length: 744802
Score: 1299.00 Matches: 253
Percent Similarity: 99.61% Conservative: 0
Best Local Similarity: 99.61% Mismatches: 1
Query Match: 17 Indels: 0
DB: 17 Gaps: 0
US-09-441-857-12 (1-254) x US-10-292-798-1369 (1-744802)
Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 345708 ATGTCGCGGGCGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTCTCTGGG 345649
Qy 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAsn 40
Db 345648 GAGCAAGCGTTGCAAGACATCTTTGATCACCAGATTCAGGTATGACGTTTTCACAAC 345589
Qy 41 ThrTyrGlnAlaIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 345588 ACCTATCAGGCAATATATGGCATTGACTTTTATCAAAAATATGTAATTTGGAGATGGA 345529
Qy 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 345528 ACAATCGGGCTTCGGCTGTGGGATACGGCGGTTCAGGAACGTCCTCCGTAGCTCATTC 345469
Qy 81 ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAsnSer 100
Db 345468 AGGTACATCCGTGATTCTGCTGCAGCTGTAGTAGTTTACCATATACAAATGTTAACTCA 345409
Qy 101 PheGlnGlnThrThrLysTyrPheAspValArgThrGluArgGlySerAspValIle 120
Db 345408 TTCAGCAAACTACAAAGTGATTCATCTCAGACACAAAGAGAGAGTGTATGTTATC 345349
Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 345348 ATCAGCTAGTAGGAAATAGAACAGATCTTGTCTGACAAGAGGCAAGTGTCTAGTTGAGGAG 345289
Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGlnThrArgAlaLysThrGly 160
Db 345288 GGAGAGAGGAAAGCCAAAGGGCTGAATGTTACGTTTATTGAAACTAGGGCAAAAGCTGGA 345229
Qy 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 345228 TACAATGTAAGCAGAGCTCTTTCCAGCTGTAGCAGCAGCTTTGCGGGAATGGAAAGCACA 345169
Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 345168 CAGGACGGAAGCAGAGAGACATGAGTGACATATAAACTGGAAGAAAGCCCTCAGGAGCAAAACA 345109
```

Qy	201	ValSerGluGlyGlyCySerCysTyrSerProMetSerSerThrLeuProGlnLys	220
Db	345108	GTCAGCGAAGGGGTGTCTGTACTCTCCCATGTCATCTTCAACGCTTCTCAGAAG	345049
Qy	221	ProProTyrSerPheIleAspCySerValAsnIleGlyLeuAsnLeuPheProSerLeu	240
Db	345048	CCCCCTTACTCTTTTCATTGATCGAGTGTAATATTGCCTTGAACCTTTTCCCTTCATT	344989
Qy	241	IleThrPheCyAsnSerSerLeuLeuProValSerTrpArg	254
Db	344988	ATAACGTTTTCGCAATTCATCATTTGCTGCTGCTCGTGGAGG	344947
RESULT 2			
US-09-925-300-424			
; Sequence 424, Application US/09925300			
; Patent No. US20020151681A1			
; GENERAL INFORMATION:			
; APPLICANT: Craig Rosen,			
; APPLICANT: Steve Ruben			
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies			
; FILE REFERENCE: PA101			
; CURRENT APPLICATION NUMBER: US/09/925,300			
; CURRENT FILING DATE: 2001-08-10			
; PRIOR APPLICATION NUMBER: PCT/US00/05988			
; PRIOR FILING DATE: 2000-03-08			
; PRIOR APPLICATION NUMBER: 60/124,270			
; PRIOR FILING DATE: 1999-03-12			
; NUMBER OF SEQ ID NOS: 1890			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 424			
; LENGTH: 3118			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (388)			
; OTHER INFORMATION: n equals a,t,g, or c			
; NAME/KEY: misc feature			
; LOCATION: (485)			
; OTHER INFORMATION: n equals a,t,g, or c			
US-09-925-300-424			
Alignment Scores:			
Pred. No.: 7.3e-134 Length: 3118			
Score: 1133.00 Matches: 224			
Percent Similarity: 89.53% Conservative: 7			
Best Local Similarity: 86.82% Mismatches: 23			
Query Match: 86.89% Indels: 4			
DB: 9 Gaps: 1			
US-09-441-857-12 (1-254) x US-09-925-300-424 (1-3118)			
Qy	1	MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly	20
Db	415	ATGTCCACGGCGGAGACTTCGGAAATCCGCTGAGGAAAATTCAGCTGGTGTCTCGGGG	474
Qy	21	GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn	40
Db	475	GAGCAAAGCKNTGGAAGACATCTTTGATCACCATTCATGTTATGACAGTTTTGACAAC	534
Qy	41	ThrTyrGlnAlaIleIleGlyLeAspPheLeuSerLysThrMetTyrLeuGluAspGly	60
Db	535	ACCTATCAGGCAACAATTTGGCATTGACITTTTATCAAAAACCTATGTTCTGGAGGATCGA	594
Qy	61	ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro	80
Db	595	ACAGTAGCATTCGAATTTATGGACACACAGCAGGTCAAGAGCGGTTTCAGGAGCTTGATTCT	654
Qy	81	ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer	100
Db	655	AGCTACATTCGTGACTCCACTGTGGCAGTTGTGTTTATGATATCACAAATGTTAACTCA	714
Qy	101	PheGlnGlnThrLysThrPileAspAspValArgThrGluArgLysSerAspValIle	120

```

Db      715  |||||TTCACGCAAACTACAAGTGGATTTGATGATGTCAGAACACAGAAAGAGGAAGTGTATGTTATC 774
Qy      121  |||||IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db      775  |||||ATCATGCTAGTAGGAATAAACAAGATCTTGTCTGACACAGAGCGAAGTGTCAATTGAGGAG 834
Qy      141  |||||GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db      835  |||||GGAGAGAGGAAAGCCAAAGAGCTGAATGTTATGTTTATTGAAACTAGTGCAAAAGCTGGA 894
Qy      161  |||||TyrAsnValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr 180
Db      895  |||||TACAATGTAAAGCAGCTCTTTTCACGCTGTAGCAGCAGCTTTGCCGGGAATGGAAAGCACA 954
Qy      181  |||||GlnAspGlySerArgGluAspMetSerAspIleLysLeuGlnLysProGlnGluGlnThr 200
Db      955  |||||CAGCACAGACGACAGAGAGATATGATTGACATAAACTCGAANAAGCCTCAGGAGCAACCA 1014
Qy      201  |||||ValSerGluGlyGlyCysSerCysThrSerProMetSerSerSerThrLeuProGlnLys 220
Db      1015  |||||GTCAGTGAAGGAGGCTGTTCTCCTGCTAATCTCCCATGTCATCTTTCAACCTTTCTTCAGAAGC 1074
Qy      221  |||||-----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
Db      1075  |||||TCACGTGTTGGCCCCCTTACTCTTTTCATTGACTGCAGTGTGAATATTGGCTTTGAACCTT 1134
Qy      237  |||||PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db      1135  |||||TTCCCTTCAGTAATAACGATTATTCGAATTCATCATCTGCTGCTCTCGTGGAGA 1188

RESULT 3
US-10-116-275-157
; Sequence 157, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Meth
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-275-157

Alignment Scores:
Pred. No.: 1,62e-112 Length: 740
Score: 959.00 Matches: 191
Percent Similarity: 90.28% Conservative: 4
Best Local Similarity: 88.43% Mismatches: 21
Query Match: 73.54% Indels: 0
DB: 17 Gaps: 0

US-09-441-857-12 (1-254) x US-10-116-275-157 (1-740)

Qy      1  MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db      71  ATGTCCACGGCGGAGACTTCGGGAATCCGCTCAGGAAATTCAAAGCTGGTGTCTCTGGGG 130

Qy      21  GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40
Db      131  GAGCAAGCCCTCGGAAGACATCTTTGATCACAGATTTATGTATGACAGCTTTTGACAAC 190

```

US-09-441-857-12 (1-254) x US-10-116-275-157 (1-740)

Alignment Scores:		
Pred. No.:	1.62e-112	Length:
Score:	959.00	Matches:
Percent Similarity:	90.28%	Conservative:
Best Local Similarity:	88.43%	Mismatches:
Query Match:	73.54%	Indels:
DB:	17	Gaps:
		0
		0
		740

; LENGTH: 740
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-116-275-157

Alignment Scores:	
Pred. No.:	7.3e-134
Score:	1133.00
Percent Similarity:	89.53%
Best Local Similarity:	86.82%
Query Match:	86.89%
DB:	9
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:
	3118
	224
	7
	23
	4
	1

US-09-441-857-12 (1-254) x US-09-925-300-424 (1-3118)

Qy	1	MetSerAlaGlyClyAspPheGlyAsnProLeuArglysPheLysLeuValPheLeuGly	20
Db	415	ATGTCCACGGCGGAGACTTCGGGAAATCCGCTGAGGAAATCAAGCTGTGTCTCTGGGG	474
Qy	21	GluGlnSerValAlaIySthrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn	40
Db	475	GAGCAAGCWNVTGGAAAGACATCTTTTGATCACCAGATTCATGTATGACGTTTGTGACAC	534
Qy	41	ThrTyrGlnAlaIleIleGlyIleAspPheLeuSerIysThrMetTyrLeuGluAspGly	60
Db	535	ACCTATCAGCGCAACAATTTGGCAATTGCATTTTATCAAAACATATGTACTTCTGGAGGATCGA	594
Qy	61	ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro	80
Db	595	ACAGTACGATTGCAATTATGGACACACAGCGGTCAAGACGGTTCAGAGCGTTGATTCTCT	654
Qy	81	ArgTyrIleArgAspSerAlaIaIaValValValTyrAspIleThrAsnValAsnSer	100
Db	655	AGCTACATTTCTGTACTCCACTGTGCGAGTGTGTGTATGTATATACAAAATGTTTAACATCA	714
Qy	101	PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle	120

```
QY 41 ThrTyrGlnAlaIlelleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 191 ACCTATCAGGCAACAAATTTGGCATTCATTTATCAAAACCTATGACTTGGAGATCGA 250
QY 61 ThrIleGlyLeuArgLeuTyrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 251 ACAGTACGATTTCAATTTATGGGACACAGCAGGTCAAGAGCGGTTTCAGGAGCTTGATTCT 310
QY 81 ArgTyrIleArgAspSerAlaAlaValValValValTyrAspIleThrAsnValAsnSer 100
Db 311 AGCTACATTCGAGCTCCACTGTGGCAGTGTGTGTATGATATCACAAATGTTAACTCA 370
QY 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 371 TTCAGCAAACTACAAAGTGGATTGATGATGTCAGAAACAGAGAGGAGTGTATTC 430
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 431 ATCATGCTAGTAGGAATAAAACAGATCTTGCTGACAAAGAGCAAGTGTCAATTGAGGAG 490
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 491 GGAGAGAGGAAGCCAAAGAGCTGAATGTTATGTTATTTGAAACTAGTGCAAAAGCTGGA 550
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 551 TACATGTAAAGCAGCTTTTCGACGTGTAGCAGCAGCTTTGCCGGGAATGGAAGCACA 610
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 611 CAGGACAGAAGCAGAGAAGATGATGATGACATATAAACTGGAAAAAGCCTCAGGAGCAACA 670
QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThr 216
Db 671 GTCAGTGAAGGAGGCTGTCTGCTTAATGTCCTAGTCATCTTCAACC 718
```

RESULT 4

```
US-10-172-118-1656
; Sequence 1656, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1656
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_016577
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1656
```

```
Alignment Scores:
Pred. No.: 1,45e-103 Length: 1266
Score: 892.00 Matches: 172
Percent Similarity: 90.38% Conservative: 16
Best local Similarity: 82.69% Mismatches: 20
Query Match: 68.40% Indels: 0
Db: 17 Gaps: 0
```

```
US-09-441-857-12 (1-254) x US-10-172-118-1656 (1-1266)
QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheIleGly 20
Db 338 ATGTCGCCAGGGGAGATTTTGGGAATCCACTGAGAAAAATTCAGAGTTGGTGTCTTGGGG 397
QY 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn 40
Db 398 GAGCAGAGCGTCGGAGAGAGCGTCTCTGATTACGAGGTTTCATGTACGACAGCTTCGACAC 457
QY 41 ThrTyrGlnAlaIlelleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 458 ACATACCAGGCAACATTTGGGATTGACTTCTTGTCAAAAACCATGTACTTGGAGGACCGC 517
QY 61 ThrIleGlyLeuArgLeuTyrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 518 ACGTGGCGACTGCGAGCTCTGGGACACAGCTGCTCAGGAGAGGTTCGCGAGCTGATCCCC 577
QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
Db 578 AGCTACATTCGGGACTCCACGGTGGCTGTGGTGTGTACGACATCACAAATCTCACTCC 637
QY 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 638 TTCCAACAGACCTCTAAGTGGATCGACGCTCAGGACAGAGAGGGCAGTGTATTATC 697
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 698 ATCATGCTGTGGGCCAACAAAGACGACCTGGCTGATTAAGAGGCAGATAACCATCGAGGAG 757
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 758 GGGAGAGCGCGCCCAAGAACTGAGCGTGTATTTGAGACCACTGCGAAAGCTGGC 817
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 818 TACAACGTGAAGCAGCTTTTTCGACGCTGTGGCTGCTCTACCCGGAATGGAGAAATGTC 877
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 878 CAGGAGAAAAACAAAGAGGGATGATCGACATCAAGCTGGACAAACCCAGGAGCCCCCG 937
QY 201 ValSerGluGlyGlyCysSerCys 208
Db 938 GCCAGCGAGGGCGGCTGCTCTCTCTC 961
```

RESULT 5

```
US-10-342-887-1656
; Sequence 1656, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1656
; LENGTH: 1266
; TYPE: DNA
```


! ORGANISM: Homo sapiens
US-10-342-887-1656

Alignment Scores:
Pred. No.: 1,45e-103 Length: 1266
Score: 892.00 Matches: 172
Percent Similarity: 90.38% Conservative: 16
Best Local Similarity: 82.69% Mismatches: 20
Query Match: 68.40% Indels: 0
DB: 17 Gaps: 0

US-09-441-857-12 (1-254) x US-10-342-887-1656 (1-1266)

```
QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
DB 338 ATGTCGCGAGGGGAGATTTGGGAATCCACTGAGAAAATTCAGTTGGTGTCTTGGGG 397
QY 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAen 40
DB 398 GAGCAGAGCGTCGGGAAGACGCTCTGTATTACGAGGTTTCATGTACGACAGCTTCGACAA 457
QY 41 ThrTyrGlnAlaLysThrSerLeuIleGlyLeuAspPheLysThrMetTyrLeuGluAspGly 60
DB 458 ACATACCGAGCAACCATTTGGGATTGACTTCTGTCAAAAACCATGTACTTGGAGGACCGC 517
QY 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
DB 518 ACGTTCGAGCTGCAGCTCTGGGACACAGCTGTGTACGAGAGGTTCCGACAGCTGTATCCC 577
QY 81 ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAenSer 100
DB 578 AGCTACATCCGGGACTCCACGGTGGCTGTGGTGTACGACATCACAAATCTCAACTCC 637
QY 101 PheGlnGlnThrThrLysTrpIleAspValArgThrGluArgGlySerAspValIle 120
DB 638 TTCCAACAGACCTCTAAAGTGGATTCGACGCTCAGGACAGAGGGGCGAGTGTATTC 697
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
DB 698 ATCATCTGTGGGCAACAGCGACTGGCTGTATAGAGCGACAGATTAACCATCGAGGAG 757
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
DB 758 GGGGAGCAGCGCCAAAGAACTGAGCGTCATGTTTCATTGAGACAGTCCGAAAGACTGCG 817
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr 180
DB 818 TACAACGTGAAGCAGCTTTTCGACGTGTGGCGTGGCTCTACCCGGAATGGAGAATGTC 877
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluInThr 200
DB 878 CAGGAGAAAAGCAAGAGGGATGATCGACATCAAGCTGGACAAACCCAGGAGCCCCCG 937
QY 201 ValSerGluGlyCysSerCys 208
DB 938 GCCAGCGAGGGCGGCTGCTCCTGC 961
```

RESULT 6

US-10-848-755A-181
; Sequence 181, Application US/10848755A
; Publication No. US20050054826A1

GENERAL INFORMATION:

APPLICANT: Mao, Mao

TITLE OF INVENTION: HUMAN DIAPHANOUS-3 GENE AND METHODS OF USE THEREFOR

FILE REFERENCE: 9301-196-999

CURRENT APPLICATION NUMBER: US/10/848,755A

CURRENT FILING DATE: 2004-05-18

PRIOR APPLICATION NUMBER: 60/471,842

PRIOR FILING DATE: 2003-05-11

NUMBER OF SEQ ID NOS: 275

SOFTWARE: PatentIn version 3.2 CAM: 301891-999188

SEQ ID NO 181

LENGTH: 1266

! TYPE: DNA
! ORGANISM: Homo sapiens
US-10-848-755A-181

Alignment Scores:
Pred. No.: 1,45e-103 Length: 1266
Score: 892.00 Matches: 172
Percent Similarity: 90.38% Conservative: 16
Best Local Similarity: 82.69% Mismatches: 20
Query Match: 68.40% Indels: 0
DB: 19 Gaps: 0

US-09-441-857-12 (1-254) x US-10-848-755A-181 (1-1266)

```
QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
DB 338 ATGTCGCGAGGGGAGATTTGGGAATCCACTGAGAAAATTCAGTTGGTGTCTTGGGG 397
QY 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAen 40
DB 398 GAGCAGAGCGTCGGGAAGACGCTCTGTATTACGAGGTTTCATGTACGACAGCTTCGACAA 457
QY 41 ThrTyrGlnAlaLysThrSerLeuIleGlyLeuAspPheLysThrMetTyrLeuGluAspGly 60
DB 458 ACATACCGAGCAACCATTTGGGATTGACTTCTGTCAAAAACCATGTACTTGGAGGACCGC 517
QY 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
DB 518 ACGTTCGAGCTGCAGCTCTGGGACACAGCTGTGTACGAGAGGTTCCGACAGCTGTATCCC 577
QY 81 ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAenSer 100
DB 578 AGCTACATCCGGGACTCCACGGTGGCTGTGGTGTACGACATCACAAATCTCAACTCC 637
QY 101 PheGlnGlnThrThrLysTrpIleAspValArgThrGluArgGlySerAspValIle 120
DB 638 TTCCAACAGACCTCTAAAGTGGATTCGACGCTCAGGACAGAGGGGCGAGTGTATTC 697
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
DB 698 ATCATCTGTGGGCAACAGCGACTGGCTGTATAGAGCGACAGATTAACCATCGAGGAG 757
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
DB 758 GGGGAGCAGCGCCAAAGAACTGAGCGTCATGTTTCATTGAGACAGTCCGAAAGACTGCG 817
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr 180
DB 818 TACAACGTGAAGCAGCTTTTCGACGTGTGGCGTGGCTCTACCCGGAATGGAGAATGTC 877
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluInThr 200
DB 878 CAGGAGAAAAGCAAGAGGGATGATCGACATCAAGCTGGACAAACCCAGGAGCCCCCG 937
QY 201 ValSerGluGlyCysSerCys 208
DB 938 GCCAGCGAGGGCGGCTGCTCCTGC 961
```

RESULT 7

US-10-108-605-44
; Sequence 44, Application US/10108605
; Publication No. US20020160934A1

GENERAL INFORMATION:

APPLICANT: Broadus, Julie

APPLICANT: Stam, Lynn

APPLICANT: Bachmann, Jane

APPLICANT: Kamdar, Kim

TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE

TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF

FILE REFERENCE: 31133B

CURRENT APPLICATION NUMBER: US/10/108,605

CURRENT FILING DATE: 2002-03-27

PRIOR APPLICATION NUMBER: US 09/761,142

Qy	42	TyrGlnAlaIleIleGlyIleAspPheLeuSerIysThrMetTyrLeuGluAspGlyThr	61
Db	366	TATCAACGCGACCATCGGAATCGACTTTCTGTCAAAAGACTATGTACTTAGAGACAGGACT	425
Qy	62	IleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIleProArg	81
Db	426	GTCCGACTCCAGCTTTGGGATACAGCCGGTCAAGAAAGATTTCAGGTCTGTTGATTCCTTCC	485
Qy	82	TyrIleArgAspSerAlaAlaAlaValValTyrAspIleThrAsnValAsnSerPhe	101
Db	486	TACATTCCGGGATTCGAGCGTTGCTGTGTGTTTACGATATTTCAACGCCCAAGTCCCTTC	545
Qy	102	GlnGlnThrThrIysTrpIleAspAspValArgThrGluArgGlySerAspValIleIle	121
Db	546	CAGAATACCCGGAATGGATTGACGCGTACGAGGGGACGTGGCAATGATGTTATCATT	605
Qy	122	ThrLeuValGlyAsnArgThrAspLeuAlaAspIysArgGlnValSerValGluGluGly	141
Db	606	GTTCTAGTTGGCAACAAGACTGATCTTAAACCAAGCGCGAGGTCAACCCCGCGAGGT	665
Qy	142	GluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGlyTyr	161
Db	666	GAACAGAGCGCTAAGAAGACCGGTTCGATGTTTTCGAGACTTAGCGCCAAAGTGTGTCAT	725

3

```

Qy 162 AsnVallysGlnLeuPheArgArgValaLaLaLeuProGlyMetGluSerThrGln 181
Db 726 AACGTCAGCAACTTTTCCGCGGATAGCCCAAGCTCTACCAGGCGATGGAGCGCAAGGT 785
Qy 182 AspGlySerArgGluAspMetSerAspIleIysLeuGluLysProGlnGlnThrVal 201
Db 786 AGTANAGGAGAGACCGAAGTGATCGATGTGAACATC-----AACCCAAAGAGACCAACCC 842
Qy 202 SerGluGlyGlyCysSerCys 208
Db 843 AACGAT---GGATCGCGCATGC 860

RESULT 10
US-10-437-963-63907/c
; Sequence 63907, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION: Thomas J.
; APPLICANT: La Rosa, David K.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

```

```

; SEQ ID NO 63907
; LENGTH: 1175
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_65103C.1
US-10-437-963-63907

```

Alignment

Spec. No.:	2,848-90	Length:	1175
Prod.	710.00	Matches:	139
Percent Similarity:	75.69%	Conservative:	26
Best Local Similarity:	63.76%	Mismatches:	47
Query Match:	54.45%	Indels:	6
DB:	18	Gaps:	2
US-09-441-857-12 (1-254) x US-10-437-963-63907 (1-1175)			

```
QY 11 LeuArgLysPheLysLeuValPheLeuGlyGluGlnSerValAlaLysThrSerLeuIle 30
Db 1065 CTCGCCAGTACAAGCTGGTGTCTCTCGCGCCAGCAGTCCGTCGCGCAAGCAGCATCATC 1006

QY 31 ThrArgPheArgTyrAspSerPheAspAenThrTyrGlnAlaIleGlyLeuAspPhe 50
Db 1005 ACCCGCTTCATGTACGACAAAGTTCGCAACACCTACAGGCTACAAATCGGTATTGACTTC 946

QY 51 LeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAla 70
Db 945 CTGTCAAAAGACAATGTACCTTCAAGATAGGACTGTACAGACTGCAACTTTGGGATACAGCT 886

QY 71 GlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaVal 90
Db 895 GGTCAAGAACAGATTTCAGAGATTGATCCAGCTATATCAGAGACTCTTCGGTTGCTGTC 826

QY 91 ValValTyrAspIleThrAsnValAsnSerPheGlnGlnThrTyrLysTyrIleAspAsp 110
Db 825 ATCGTATTGATGTTGCAAGCAGACAGCTTTCTTAACACTTCAAAATGGATAGAGAA 766

QY 111 ValArgThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeu 130
Db 765 GTTAGAACTGAGAGGGCAGTGTATCATTTGCTGTGGTGGGAAACAAACTGATCTT 706

QY 131 AlaAspLysArgGlnValSerValGluGlyGluArgLysAlaLysGlyLeuAsnVal 150
Db 705 GTTGCAAGAGCAAGTCTCCATAGAGGAGAGGCAAGCGAAAGATCTCGGTGTG 646

QY 151 ThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgVal 170
Db 645 ATGTTTCATGAAACCAAGTCTTAAGCTGGTTTAACTAATAGGCGTTGTTCCGTAAGATT 586

QY 171 AlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAspMetSerAsp 190
Db 585 GCGGTGCACCTTCCTGGAATGAGACCCCTCTCATCAGCAAAAGCAGGAGACATGGTTGAT 526

QY 191 IleLysLeuGlu-----LysProGlnGluGlnThrValSerGluGlyGlyCys 206
Db 525 GTGAACCTTGAAGTCAAGCAAGCCCACTCATCCAGTCTCAGGCACAGGCTGGGGGATGC 466

QY 207 SerCys-----TyrSerProMetSerSerThrLeuProGlnLysProPro 222
Db 465 AGTTGTTAGTTCAGTCCCTTAACTACTGCTTCGCGGTCTCAAGGCCCCCC 412
```

RESULT 11

```
US-10-425-114-4106
; Sequence 4106, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 4106
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700349145_FLI
US-10-425-114-4106
```

```
Alignment Scores:
Pred. No.: 3.36e-80 Length: 1200
Score: 709.50 Matches: 144
```

```
Percent Similarity: 72.46% Conservative: 27
Best Local Similarity: 61.02% Mismatches: 50
Query Match: 54.41% Indels: 15
DB: 17 Gaps: 3

US-09-441-857-12 (1-254) x US-10-425-114-4106 (1-1200)

QY 11 LeuArgLysPheLysLeuValPheLeuGlyGluGlnSerValAlaLysThrSerLeuIle 30
Db 181 CTCGCCAGTACAAGCTGGTGTCTCTCGCGGAGCAGTCCGTCGCGCAAGCAGCATCATC 240

QY 31 ThrArgPheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleGlyLeuAspPhe 50
Db 241 ACCCGCTTCATGTACGACAAAGTTCGATACACCTTACAGGCTACGATGGTATTGATTTC 300

QY 51 LeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAla 70
Db 301 CTGTCAAAAGACAATGTACCTTGAAGATAGAACTGTGAGACTCCCAACTCTCGGATACTGCT 360

QY 71 GlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaVal 90
Db 361 GGTCAAGAAAGATTTCAGAGTTTAAATTCGAAGCTATATTAGAGACTCTTCAGTTGCTGTC 420

QY 91 ValValTyrAspIleThrAsnValAsnSerPheGlnGlnThrTyrLysTyrIleAspAsp 110
Db 421 ATGTTATTCGATGTTGCAAGCAGCAGCTCTCTTAATAATACATCTAAGTGGATAGAGAA 480

QY 111 ValArgThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeu 130
Db 481 GTTCGCACTGAGAGGGCAGTGTATCATTTGCTGTGGGAAACAAACTGACCTT 540

QY 131 AlaAspLysArgGlnValSerValGluGlyGluArgLysAlaLysGlyLeuAsnVal 150
Db 541 GTTGCAAGAGCAAGTCTCAATAGAGGAGGAGGCAAGCGCAAGGACCTTGGAGTC 600

QY 151 ThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgVal 170
Db 601 ATGTTTATGAAACCAAGTGTCTAAGCTGGGTTTAACTAATAGGCGTGTTCGCTAAATTT 660

QY 171 AlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAspMetSerAsp 190
Db 661 GCTGTCACCTTCCTGGAATGAGACGCTCTCATCAGCAAGCAGGAGACATGGTTGAT 720

QY 191 IleLysLeuGlu-----LysProGlnGluGlnThrValSerGluGlyGlyCys 206
Db 721 GTGAACCTGAGTCCGCAATGCAAACTCGTCCAGTCTCAGGCTCAGGCTGGGGGATGC 780

QY 207 SerCysTyrSerProMetSerSerThrLeuProGlnLysProProTyrSerPheIle 226
Db 781 AGTTGT-----TAGTTGCACTCTCTGACACTTGTCT----- 810

QY 227 AspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeuIleThr 242
Db 811 ---TGCTCATGATCTTTGGGACTCTGAGCATCAGAGGCGGTATCATCT 855
```

RESULT 12

```
US-10-425-114-1068
; Sequence 1068, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1068
```

; LENGTH: 1220

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700101038_FLI

US-10-425-114-1068

Alignment Scores:

Pred. No.: 3,45e-80 Length: 1220
Score: 709.50 Matches: 144
Percent Similarity: 72.46% Conservative: 27
Best Local Similarity: 61.02% Mismatches: 50
Query Match: 54.41% Indels: 15
DB: 17 Gaps: 3

US-09-441-857-12 (1-254) x US-10-425-114-1068 (1-1220)

QY 11 LeuArgLysPheLysLeuValPheLeuGluGlnSerValAlaLysThrSerLeu 30
DB 197 CTCGCCAAGTACAAGCTCGTCTCTCTGGGGACCAAGTCCGTCGCAAGACGACATCATC 256
QY 31 ThrArgPheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGlyLeuAspPhe 50
DB 257 ACCCGCTTCATGTACGACAAAGTTCGATTAACCTTACCAGGCTACGATTGGTATTGATTTC 316
QY 51 LeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAla 70
DB 317 CTGTCAAGACAAATGACCTTCAAGATAGAACTGTGAGACTCCAACTCTGGGATACTGCT 376
QY 71 GlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaVal 90
DB 377 GGTCAAGAAAGATTCAAGGAGTTAAATTCGAAGCTATATTAGAGACTCTTCAGTTGCTGTC 436
QY 91 ValValTyrAspIleThrAsnValAsnSerPheGlnGlnThrThrLysThrIleAspAsp 110
DB 437 ATTGATTTCGATGTTGCAAGCAGGAGCTCTTCTTAAATACATCTAAGTGGATAGAGAA 496
QY 111 ValArgThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeu 130
DB 497 GTTCGACTGAGAGGGGCGAGTGATGTTATCATTTGCTTGTGGGACAAACATGACCTT 556
QY 131 AlaAspLysArgGlnValSerValGluGluGlyLysAlaLysGlyLeuAsnVal 150
DB 557 GTTGACAAGAGCAAGTCTCAATAGAGGAAGGGGCAAGCGCAAGGACCCCTGGAGTC 616
QY 151 ThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgVal 170
DB 617 ATGTTATTGAACCAAGTCTAAGGCTGGGTTTAACTTAAGCGCTGTTCCGTAAATTT 676
QY 171 AlaAlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAspMetSerAsp 190
DB 677 GCTGCTGCACCTTCCTGGAATGAGAGCGCTCTCATCAGCGAAGCAGGACATGTTGAT 736
QY 191 IleLysLeuGlu-----LysProGlnGluGlnThrValSerGluGlyCys 206
DB 737 GTGAACCTTGAGGTCGCGCAATGCAAACTCGTCCAGTCTCAGGCTCAGGCTGGGGATGC 796
QY 207 SerCysTyrSerProMetSerSerThrLeuProGlnLysProProTyrSerPheIle 226
DB 797 AGTTGT-----TAGTTGAGTCTCTGACACTTGCT----- 826
QY 227 AspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeuIleThr 242
DB 827 ---TGCTCATGATCTTTGGGACTCTGAGCATCAGAAGGCGTTACTACT 871

RESULT 13

US-10-425-115-20462

; Sequence 20462, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 20462
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1506)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MET4577_118663C.1
US-10-425-115-20462

Alignment Scores:

Pred. No.: 4.77e-80 Length: 1506
Score: 709.50 Matches: 144
Percent Similarity: 72.46% Conservative: 27
Best Local Similarity: 61.02% Mismatches: 50
Query Match: 54.41% Indels: 15
DB: 18 Gaps: 3

US-09-441-857-12 (1-254) x US-10-425-115-20462 (1-1506)

QY 11 LeuArgLysPheLysLeuValPheLeuGluGlnSerValAlaLysThrSerLeu 30
DB 335 CTCGCCAAGTACAAGCTCGTCTCTGGGGACCAAGTCCGTCGCAAGACGACATCATC 394
QY 31 ThrArgPheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGlyLeuAspPhe 50
DB 395 ACCCGCTTCATGTACGACAAAGTTCGATTAACACCTTACCAGGCTACGATTGGTATTGATTTC 454
QY 51 LeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAla 70
DB 455 CTGTCAAGACAAATGATACCTTGAAGATAGAACTGTGAGACTCCAACTCTGGGATACTGCT 514
QY 71 GlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaVal 90
DB 515 GGTCAAGAAAGATTCAAGGAGTTAAATTCGAAGCTATATTAGAGACTCTTCAGTTGCTGTC 574
QY 91 ValValTyrAspIleThrAsnValAsnSerPheGlnGlnThrThrLysThrIleAspAsp 110
DB 575 ATGTTATTGAACCAAGTCTAAGGCTGGGTTTAACTTAAGCGCTGTTCCGTAAATTT 634
QY 111 ValArgThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeu 130
DB 635 GTTCGACTGAGAGGGGCGAGTGATGTTATCATTTGCTTGTGGGACAAACATGACCTT 694
QY 131 AlaAspLysArgGlnValSerValGluGluGlyLysAlaLysGlyLeuAsnVal 150
DB 695 GTTGACAAGAGCAAGTCTCAATAGAGGAAGGGGCAAGCGCAAGGACCCCTGGAGTC 754
QY 151 ThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgVal 170
DB 755 ATGTTATTGAACCAAGTCTAAGGCTGGGTTTAACTTAAGCGCTGTTCCGTAAATTT 814
QY 171 AlaAlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAspMetSerAsp 190
DB 815 GCTGCTGCACCTTCCTGGAATGAGAGCGCTCTCATCAGCGAAGCAGGACATGTTGAT 874
QY 191 IleLysLeuGlu-----LysProGlnGluGlnThrValSerGluGlyCys 206
DB 875 GTGAACCTTGAGGTCGCGCAATGCAAACTCGTCCAGTCTCAGGCTCAGGCTGGGGATGC 934
QY 207 SerCysTyrSerProMetSerSerThrLeuProGlnLysProProTyrSerPheIle 226
DB 935 AGTTGT-----TAGTTGAGTCTCTGACACTTGCT----- 964

```
Qy 227 AspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeuIleThr 242
Db 965 ---TGCTCATGATCTTTGGGACTCTGAGCATCAGAGCGGTATCACT 1009

RESULT 14
US-10-425-115-20463
; Sequence 20463, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 20463
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_118664C.1
US-10-425-115-20463

Alignment Scores:
Pred. No.: 7,41e-80 Length: 1242
Score: 707.00 Matches: 136
Percent Similarity: 78.22% Conservative: 22
Best Local Similarity: 67.33% Mismatches: 40
Query Match: 54.22% Indels: 4
DB: 18 Gaps: 1

US-09-441-857-12 (1-254) x US-10-425-115-20463 (1-1242)

Qy 11 LeuArgLysPheLysLeuValPheLeuGlyGlnSerValAlaLysThrSerLeuIle 30
Db 193 CTCGCCAAGTACAAAGCTCGTCTCTCGGGGACAGCTCCGCGGCAAGCAGCATCATC 252

Qy 31 ThrArgPheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGlyLeuAspPhe 50
Db 253 ACCCGCTTCATGATAGTAAGTTCGACCAACTTACCAGGCTACAAATGGTATTGATTTC 312

Qy 51 LeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAla 70
Db 313 CTGTCAAAGACAAATGTACTTGAAGATAGAACTGTGAGACTCCAACCTCTGGGATACAGCT 372

Qy 71 GlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaVal 90
Db 373 GGTGAGGAAAGGTTTCAGGAGTTTAATTCCAAGCTATATACAGAGACTCTTCAGTTGCTGTC 432

Qy 91 ValValTyrAspIleThrAsnValAsnSerPheGlnGlnThrLysThrLysTrpIleAspAsp 110
Db 433 ATTGTATTCGAITGTCGAGCAGGAGCTCTTCTTAATATCCTCAAGTGGATAGATGAA 492

Qy 111 ValArgThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeu 130
Db 493 GTTCGCACTGAGAGGGGCGAGTGATGTTATCTGCTTGTGGGACAAACAACTGACCTT 552

Qy 131 AlaAspLysArgGlnValSerValGluGlyGlnArgLysAlaLysGlyLeuAsnVal 150
Db 553 GTTGACAAAGGCGCAAGTCTCGATAGAGGAGGAGGCAAGGCAAGGCAAGCTTGGTGTGTC 612

Qy 151 ThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgVal 170
Db 613 ATGTTTATGAAACACAGTCTTAAGCTGGGTTTAACATTAAAGGCGCTGTTCCGTAAATTT 672

Qy 171 AlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAspMetSerAsp 190
Db 673 GCTGCTCCCTTCTCGGTATGGAGACACTCTCATCAGCGAAGCAGGAAGCATGGTTGAT 732
```

```
Qy 191 IleLysLeuGlu-----LysProGlnGlnThrValSerGluGlyGlyCys 206
Db 733 GTGAACCTTGAGGTCTGGCAATGCAAACTCGTCCAGTCTCAGGCTCAGGCTGGGGGATGC 792

Qy 207 SerCys 208
Db 793 AGTTGT 798

RESULT 15
US-10-425-114-27873
; Sequence 27873, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27873
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4741-002-F12_FLI
US-10-425-114-27873

Alignment Scores:
Pred. No.: 1.25e-79 Length: 1308
Score: 705.50 Matches: 143
Percent Similarity: 72.46% Conservative: 28
Best Local Similarity: 60.59% Mismatches: 50
Query Match: 54.10% Indels: 15
DB: 17 Gaps: 3

US-09-441-857-12 (1-254) x US-10-425-114-27873 (1-1308)

Qy 11 LeuArgLysPheLysLeuValPheLeuGlyGlnSerValAlaLysThrSerLeuIle 30
Db 281 CTCGCCAAGTACAAAGCTCGTCTCTCGGGGACAGCTCCGCGGCAAGCAGCATCATC 340

Qy 31 ThrArgPheArgTyrAspSerPheAspAsnThrTyrGlnAlaIleIleGlyLeuAspPhe 50
Db 341 ACCCGCTTCATGATAGTAAGTTCGACCAAGTTCGATACACCTACCAGGCTAGGATTGGTATTGATTTC 400

Qy 51 LeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLeuArgLeuTyrAspThrAla 70
Db 401 CTGTCAAAGACAAATGTACTTGAAGATAGAACTGTGAGACTCCAACCTCTGGGATACCTGCT 460

Qy 71 GlyGlnGluArgLeuArgSerLeuIleProArgTyrIleArgAspSerAlaAlaVal 90
Db 461 GGTGCGGAAAGAGTTCAGGAGTTTAATTCCAAGCTATATATTAGAGACTCTTCAGTTGCTGTC 520

Qy 91 ValValTyrAspIleThrAsnValAsnSerPheGlnGlnThrLysTrpIleAspAsp 110
Db 521 ATTGTATTCGATGTTGCAAGCAGGAGCTCTTCTTAATATCCTCAAGTGGATAGAGAA 580

Qy 111 ValArgThrGluArgGlySerAspValIleIleThrLeuValGlyAsnArgThrAspLeu 130
Db 581 GTTCGCACTGAGAGGGGCGAGTGATGTTATCTGCTTGTGGGACAAACAACTGACCTT 640

Qy 131 AlaAspLysArgGlnValSerValGluGlyGlnArgLysAlaLysGlyLeuAsnVal 150
Db 641 GTTGACAAAGGCGCAAGTCTCAATAGAGGAGGAGGCAAGGCAAGGCAAGCTTGGAGTC 700

Qy 151 ThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgVal 170
Db 151 ThrPheIleGluThrArgAlaLysThrGlyTyrAsnValLysGlnLeuPheArgVal 170
```

```

Db      701 ATGTTATTGAACCAAGTCTAAGCTGGGTTTAAACATTAAAGCGCTGTTCGTAAAAATT 760
Qy      171 AlaAlaLeuProGlyMetGluSerThrGlnAspGlySerArgGluAspMetSerAsp 190
      |||||
Db      761 GCTGTGCACCTTCCTGGAAATGGAGAGCTCTCATCAGCGAAACAGAGACATGGTTGAT 820
      |||||
Qy      191 IleLysLeuGlu-----LysProGlnGluGlnThrValSerGluGlyGlyCys 206
      ::|||
Db      821 GTGAACCTTGAGGTCGCGCAATGCAAACTCGTCCAGTCTCAGGCTCAGGCTGGGGGATGC 880
      ::|||
Qy      207 SerCysTyrSerProMetSerSerThrLeuProGlnLysProTyrSerPheIle 226
      |||||
Db      881 AGTTGT-----TAGTTGCAGTCTCTGACACTTGCT----- 910
      |||||
Qy      227 AspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeuIleThr 242
      |||||
Db      911 ---TGCTCATGATCTTTGGGACTCTGAGCATCAGAGGCGTTATCACT 955
      |||||

```

Search completed: April 25, 2005, 09:38:56
 Job time : 6909 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2005, 02:37:36 ; Search time 3095 Seconds
(without alignments)
3123.853 Million cell updates/sec

Title: US-09-441-857-12

Perfect score: 1304

Sequence: 1 MSAGDGFNGPLKPKLVFLG.....NLFPSLITFCNSLLLPVSWR 254

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q/cn2_1/USPTO_spool_p/US09441857/runat_22042005_155509_24759/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09441857 @CN 1 1 5180 @runat_22042005_155509_24759 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1304	100.0	3084	3 HSM801695	AL136727 Homo sapi
2	1128	86.5	1928	3 AK084131	AK084131 Mus muscu
3	1128	86.5	1955	3 AK051246	AK051246 Mus muscu
4	1128	86.5	3404	3 AK083262	AK083262 Mus muscu
5	1081	82.9	790	4 BI694279	BI694279 603347724
6	1065	81.7	897	5 BQ422263	BQ422263 AGENCOURT
7	1016	77.9	744	6 CB320435	CB320435 AGENCOURT
8	962	73.8	809	7 CN227000	CN227000 RJB008D12
9	956.5	73.4	797	7 CF151490	CF151490 AGENCOURT

10	953	73.1	792	6	CA316338	UI-M-FW0-
11	944	72.4	628	6	CB565915	CB565915 AGENCOURT
12	941.5	72.2	870	6	CA471443	CA471443 AGENCOURT
13	936	71.8	615	6	CB067426	CB067426 1q36f10.X
14	936	71.8	788	6	CB696452	CB696452 2F101-P00
15	934	71.6	696	5	BX852687	BX852687 BX852687
16	928	71.2	656	1	A1435940	A1435940 th80c04.X
17	927.5	71.1	803	5	BU379426	BU379426 603811211
18	925	70.9	866	1	AL875722	AL875722 AL875722
19	924	70.9	866	5	BQ632282	BQ632282 1l25a01.X
20	923	70.8	838	7	CK315856	CK315856 SB02028A1
21	921	70.6	858	7	CF547798	CF547798 AGENCOURT
22	919	70.5	878	6	CD361397	CD361397 AGENCOURT
23	915	70.2	738	6	CA372914	CA372914 646911 NC
24	911	69.9	804	7	CK302212	CK302212 SB02015A1
25	909.5	69.7	759	5	BP709291	BP709291 BP709291
26	907	69.6	599	7	CV025778	CV025778 3550 Full
27	905.5	69.4	914	6	CA974052	CA974052 AGENCOURT
28	905	69.4	1864	3	CR692686	CR692686 Tetraodon
29	904	69.3	764	5	BU358285	BU358285 603476202
30	901.5	69.1	876	5	BQ732681	BQ732681 AGENCOURT
31	900	69.0	762	2	BE965054	BE965054 601658839
32	900	69.0	776	7	CO421669	CO421669 GGEZHT102
33	899	68.9	646	7	CV028279	CV028279 6756 Full
34	897.5	68.8	2563	3	AK035893	AK035893 Mus muscu
35	896	68.7	757	7	CK302742	CK302742 SB02016A2
36	894.5	68.6	715	4	BG016924	BG016924 df67d09.X
37	894	68.6	871	5	BU381224	BU381224 603857076
38	893.5	68.5	833	7	CK352986	CK352986 AGENCOURT
39	891.5	68.4	854	5	BX734058	BX734058 BX734058
40	890	68.3	798	7	CF290541	CF290541 AGENCOURT
41	886	67.9	767	4	BI754357	BI754357 603026905
42	885	67.9	632	2	BF431157	BF431157 7008C05.X
43	882	67.6	792	7	CN081076	CN081076 EC28BA18D
44	881.5	67.6	886	5	BU363505	BU363505 603785110
45	879	67.4	643	2	BB628557	BB628557 BB628557

ALIGNMENTS

RESULT 1
HSM801695
LOCUS Homo sapiens mRNA; cDNA DKFZp566K144 (from clone DKFZp566K144). linear HTC 22-SEP-2004
DEFINITION AL136727
ACCESSION AL136727.1 GI:12052972
VERSION
KEYWORDS HTC.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3084)
AUTHORS Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
The German cDNA Consortium
CONSTRM Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp566K144) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp566K144
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1..3084
/organism="Homo sapiens"
/mol_type="mRNA"

```

/db_xref="RZPD:DKFZp566K144"
/db_xref="taxon:9606"
/clone="DKFZp566K144"
/tissue_type="kidney"
/clone_lib="566 (synonym: hfkd2). Vector pAMP1; host
xl-2blue; sites NotI + SalI"
/dev_stages="fetal"
/notes="RAB6C, member RAS oncogene family"
1..3084
/gene="DKFZp566K144"
456..1220
/gene="DKFZp566K144"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB6661.1"
/db_xref="GI:12052973"
/db_xref="GOA:Q9H0N0"
/db_xref="UniProt/Swiss-Prot:Q9H0N0"
/translation="MSAGSDFGNPLRKFVFLCEQSVAKTSLITRPFYDSFDNTYQA
IIGIDFLSKTWLEDTGIGRLWDTAGQERLSLIPYIRDSAAAVVVDITNVNSFQ
QTWIDVDRTERGSDVITLVGNKTDIADKRQVSEGERKAKGLNVFTETRAKGT
YNNQLFRVAALPMESTQDGSREDMSDIKLEKPOQTIVSEGGSCYSPMSSTLP
QKPPYSFIDCSVNIGLNLFPSLITFCNSLLPVSWR"
ORIGIN
Alignment Scores:
Pred. No.: 5,696-135 Length: 3084
Score: 1304.00 Matches: 254
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 3 Gaps: 0

US-09-441-857-12 (1-254) x HSM801695 (1-3084)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgGlyLysPheLysLeuValPheLeuGly 20
|||||
Db 456 ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGGAATTCAGCTGGTGTCTCTGGGG 515
|||||

Qy 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgGlyTyrAspSerPheAspAsn 40
|||||
Db 516 GAGCAAAAGCGTTGCAAGACATCTTTGATCACCAGATTGAGTATGACAGTTTGGACAA 575
|||||

Qy 41 ThrTyrGlnAlaIleIleGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
|||||
Db 576 ACCTATCAGGCAATAAATGGCATTGACTTTTATCAAAACATGATGATCTGGAGATGGA 635
|||||

Qy 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
|||||
Db 636 ACNATCGGGCTTCGGCTGGGATACGGGGGTGAGAACGCTCTCCGTAGCTCATTCCTCC 695
|||||

Qy 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
|||||
Db 696 AGGTACATCCGTGATTCTGCTGCAGCTGATGATGATGATGATGATGATGATGATGATG 755
|||||

Qy 101 PheGlnGlnThrThrLysThrIleAspAspValArgThrGluArgGlySerAspValIle 120
|||||
Db 756 TTCCAGCAAACTACAAAGTGGATTGATGATGATGATGATGATGATGATGATGATGATG 815
|||||

Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
|||||
Db 816 ATCAGCTAGTAGGAATAGAACAGATCTTCTGACAGAGCAAGTGTGAGTGGAGAG 875
|||||

Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
|||||
Db 876 GGAGAGAGCAAGCAAGGGGCTGAATGTTACGTTTATTGAAACTAGGGCAAAACTGGA 935
|||||

Qy 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
|||||
Db 936 TACAATGTAAGACGAGCTCTTTTCGACGTGTAGCAGCAGCTTTGGCCGGGAATGGAAGACA 995
|||||

Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
|||||

```

```

Db 996 CAGGCGGAAGCAGAGAGACATGATGACATATAAACTGGAAAGCCTCAGGAGCAACA 1055
Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerThrLeuProGlnLys 220
|||||
Db 1056 GTACAGCAAGGGGTGTTCCTGCTACTCTCCCATGTCTTCAACCCCTTCTCAGAG 1115
Qy 221 ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLeu 240
|||||
Db 1116 CCCCTTACTCTTTTCATTCATGTCAGTGTGAATATTGGCTTGAACCTTTTCCCTTCATTA 1175
Qy 241 IleThrPheCysAsnSerSerLeuLeuProValSerTyrArg 254
|||||
Db 1176 ATACGTTTTCATTCATTCATTCCTGCTGCTCTCTGAGG 1217

RESULT 2
AK084131
LOCUS
DEFINITION
AK084131 1928 bp mRNA linear HTC 03-APR-2004
Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
enriched library, clone:DI30097N11 product:RAB6, member RAS
oncogene family, full insert sequence.
ACCESSION
AK084131.1 GI:26350968
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
2
JOURNAL
MEDLINE
PUBMED
99279253
10349636
3
REFERENCE
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
4
JOURNAL
MEDLINE
PUBMED
20530913
11076861
5
REFERENCE
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
6
JOURNAL
MEDLINE
PUBMED
420, 563-573 (2002)
7
REFERENCE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
8
JOURNAL
MEDLINE
PUBMED
420, 563-573 (2002)
9
REFERENCE
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,

```

Okazaki, Y., Saitoh, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

FEATURES

source

1..1928
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:DL130097N11"
/db_xref="taxon:10090"
/clone="DL130097N11"
/tissue_type="spinal ganglion"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
447..1073
/note="unnamed protein product: RAB6, member RAS oncogene family (MGD|MG1:894313, GB|NM_024287, evidence: BLASTN, 99%, match=1919)
putative"
/codon_start=1
/protein_id="BAC39121.1"
/db_xref="GI:26350969"
/translation="MSAGDGFNPLRKLVLPLGQSGVSKTSLTRFMYDSFDNTYQA
TIGIDFLSKMYLEDTVRLQLWDTAGQERFRSLIPSRDSTVAVVYDITNVSPQ
QTTKWIDVTRGSDVIMLVGNKTDIADKRFQVIEGERKAKELNMFETISAKAG
YNVKQLFRVAALPGMESTQDRSDREDMIDIKLEKPPQPVNEGSC"

CDS

1073..1928
/note="unnamed protein product: RAB6, member RAS oncogene family (MGD|MG1:894313, GB|NM_024287, evidence: BLASTN, 99%, match=1919)
putative"
/codon_start=1
/protein_id="BAC39121.1"
/db_xref="GI:26350969"
/translation="MSAGDGFNPLRKLVLPLGQSGVSKTSLTRFMYDSFDNTYQA
TIGIDFLSKMYLEDTVRLQLWDTAGQERFRSLIPSRDSTVAVVYDITNVSPQ
QTTKWIDVTRGSDVIMLVGNKTDIADKRFQVIEGERKAKELNMFETISAKAG
YNVKQLFRVAALPGMESTQDRSDREDMIDIKLEKPPQPVNEGSC"

ORIGIN

Alignment Scores:
Pred. No.: 1
Score: 1128.00 Length: 1928
Matches: 224
Percent Similarity: 89.53% Conservative: 7
Best Local Similarity: 86.82% Mismatches: 23
Query Match: 86.50% Indels: 4
Gaps: 1

US-09-441-857-12 (1-254) x AK084131 (1-1928)

Qy 1 MetSerAlaGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 447 ATGTCGGCGGGCGAGACTTCGGGAATCCGCTGAGAAATTCAGCTGTGTTCTGGGA 506
Qy 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAspAsn 40
Db 507 GAGCAGAGCGTTGGAAAGACGCTCTTGATCACCCTGATCATGTATGACATTTTGACAC 566
Qy 41 ThrTyrGlnAlaIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 567 ACCATCAGGCAACAATTGGCAATTCCTTATCAAAACAATGTTACTTGGAGATAGA 626
Qy 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLysSerLeuLeuPro 80
Db 627 ACCGTGCGGATTCATTTATGGGACACAGCAGGTCAAGCGGTCAGGAGCTTGATCTCT 686
Qy 81 ArgTyrIleArgAspSerAlaAlaValValTyrAspIleThrAsnValAsnSer 100

Db 687 AGCTACATTCGAGACTCCACTGTGGCAGTGTGTTTATGATATCACAAATGTTAACTCA 746
Qy 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 747 TTCGAGCAAACTACAAATGGATGATGATGTCAGAACAGAGAGGAGGATGATGTCATC 806
Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 807 ATCATGCTAGTAGGAATAAAACAGATCTTCCTGATAAGAGGCAAGTGTCAATGAGGAG 866
Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheLeuThrArgAlaLysThrGly 160
Db 867 GGAGAGAGAAAGCAAGAGAGCTGAATGTTATTTATGAACAGTGCAGGAGCAGGA 926
Qy 161 TyrAsnValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 927 TACATGTAAAGCAGCTTTCCGGCGTGTCTGCGAGCTTTACCTGGAAATGGAAGCACA 986
Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 987 CAGCAGCAAGCAGAGAGACATGATTGACATAAACTGGAAAGCCTCAGGAGCAACCA 1046
Qy 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 1047 GTCAATGAAGCGCGCTGTCTCTGCTTAATCCCTCGCATCTTTCAACCTCTCTCGCAGG 1106
Qy 221 -----ProProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
Db 1107 CTCAGCTGTTGGGCCCATCTCTTTCAATGACGTGAGTGAATATTGGCTTGAACCTT 1166
Qy 237 PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254
Db 1167 TTCCCTTCAGTAATACGTATTGCAATTCATCATTCGTCGCTGTTTCTGCGGAGG 1220
RESULT 3
AK051246
LOCUS
DEFINITION
AK051246 1955 bp mRNA linear HTC 03-APR-2004
Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:DL130023A03 product:RAB6, member RAS oncogene family, full insert sequence.
ACCESSION
VERSION AK051246.1 GI:26341819
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED 10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED 11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE PUBMED REFERENCE AUTHORS	20530913 11076861 4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1955) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saiboh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers	DB: 3 Gaps: 1 US-09-441-857-12 (1-254) x AK051246 (1-1955) QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20 Db 474 ATGTCGCGCGGCGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTCTCTCGGA 533 QY 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAspAsn 40 Db 534 GAGCAGAGCGGTGGAAAGAGCGTCTTGATCACCAGATTCATGTATGACAGATTTGACAAC 593 QY 41 ThrTyrGlnAlaLeuLeuGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60 Db 594 ACCATATCAGCAACAATGGCATTTGACTTCTTATCAAAAACAATGTACTTGGAGGATAGA 653 QY 61 ThrLeuGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLysSerLeuLeuPro 80 Db 654 ACCGTGCGATTGCAATATATGGACACAGCAGGTCAAGAGCGGTTTCAGGAGCTTGATTCCT 713 QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100 Db 714 AGCTACATTCGAGATCCACTGTGGCGAGTTGTTGTTATGATATCAAAAATGTTAACTCA 773 QY 101 PheGlnGlnThrThrLysTyrIleAspValArgThrGluArgGlySerAspValIle 120 Db 774 TTCAGCAAACTACAAAATGGATTGATGATCAGACAGAGAGGAGGAGTGTGTCATC 833 QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140 Db 834 ATCATGCTAGTAGGAAATAAAACAGATCTTGCTGATAGAGGCAAGTGTCAATTTGAGGAG 893 QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160 Db 894 GGAGAGAGGAGAGCAAGAGAGTGAATGTTATGTAACCACTGCAAAAGCAGGA 953 QY 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180 Db 954 TACATGTTAAGCAGCTTTTCCGGCGTGTGCTGCAGCTTTACTTGGAAATGGAAGACACA 1013 QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluThr 200 Db 1014 CAGGACAGAGCAGAGAGACATGATTGACATAAACTGGAAGAGCCCTCAGGACCAACCA 1073 QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220 Db 1074 GTCAATGAGGCGCGCTGTTCTCTGCTAATCCCTTGGCATCTTTCAACCCCTCTGCAAG 1133 QY 221 -----ProTyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236 Db 1134 CTCACTGCTTTGGCCCCATACTCTTTTCATTGACTGCGAGTGTGAATATGCTTGAACCTT 1193 QY 237 PheProSerLeuLeuThrPheCysAsnSerSerLeuLeuProValSerTrpArg 254 Db 1194 TTCCCTTCAGTAATAACGTATTGCAATTCATCATCTGCTGCTGTTCTGCTGGAGG 1247 RESULT 4 AK083262 3404 bp mRNA linear HTC 03-APR-2004 LOCUS Mus musculus adult male hippocampus cDNA, RIKEN full-length DEFINITION enriched library, clone: C630031P05 product: RAB6, member RAS oncogene family, full insert sequence. ACCESSION AK083262 VERSION AK083262.1 GI:26350388 KEYWORDS HTC; CAP trapper. SOURCE Mus musculus (house mouse) ORGANISM Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. REFERENCE 1 AUTHORS Carninci, P. and Hayashizaki, Y. TITLE High-efficiency full-length cDNA cloning JOURNAL Meth. Enzymol. 303, 19-44 (1999)
---	--	---

ORIGIN

Alignment Scores:

Pred. No.: 1,82e-115 Length: 1955
 Score: 1128.00 Matches: 224
 Percent Similarity: 89.53% Conservative: 7
 Best Local Similarity: 86.82% Mismatches: 23
 Query Match: 86.50% Indels: 4

CDS

FEATURES

source

1..1955
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="FANTOM,DB:D130023A03"
 /db_xref="taxon:10090"
 /clone="D130023A03"
 /tissue_type="spinal ganglion"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="12 days embryo"
 474..1100
 /note="unnamed protein product; RAB6, member RAS oncogene family (MGD|MG1:894313, GB|NM_024287, evidence: BLASTN, 99% match=1919)
 putative"
 /codon_start=1
 /protein_id="BAC34572.1"
 /db_xref="GI:26341820"
 /translation="MSAGGDFGNPLRKFPLVGLGEQSVGKTLITRFMYDSFDNTYQA
 TIGIDFLSKTWLEDRVQLQDLTAGGERPSLIPSYIRDSVAVVVVDITNVNSFQ
 QTTWKIDVRRERSDVIIMLVGNKTLADLRQVSIIEGKAKELNMFIETSAKAG
 INVKQLFRVAALPGMESTQDRSDRMIDIKLEKPFQEPVNEGCGSC"

```

MEDLINE
PUBMED
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3404)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahita, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source
Location/Qualifiers
1..3404
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C630031P05"
/db_xref="taxon:10090"
/clone="C630031P05"
/sex="male"
/tissue_type="hippocampus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
463..1089
/note="unnamed protein product; RAB6, member RAS oncogene
family (MCD|MGI:894313, GB|NM_024287, evidence: BLASTN,
99%, matches=1919)
putative"
/codon_start=1
/protein_id="BAC38834.1"
/db_xref="GI:26350389"
/translators="MSAGDFGNLRKFKLVFLGEQSVGKTSLLTRFMYDSFDNTYQA
TIGDPLSKTWLEDRVRLQNDTACGERFRSLIPSYRSTVAVVYDITNVNSFO
QTTKWDVDRTERGSDVILMLVGNKTLADKQVSIIEGERKAKELNNVFIETSAKAG
YNVQLFRVRAALPGHESTQDRSDMIDIKLEKPOBPQVNEGSC"
3379..3384
polyA_signal
/note="putative"
3404
polyA_site
/note="putative"
ORIGIN
Alignment Scores:
Pred. No.: 4,11e-115 Length: 3404
Score: 1128.00 Matches: 224
Percent Similarity: 89.53% Conservative: 7
Best Local Similarity: 86.82% Mismatches: 23
Query Match: 86.50% Indels: 4
DB: 3 Gaps: 1
US-09-441-857-12 (1-254) x AK083262 (1-3404)
QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
Db 463 ATGTCGCGGGCGGAGACTTCGGGAATCCGCTGAGGAAATCAAGCTGGTGTCTCTGGGA 522
QY 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAsp 40
Db 523 GAGCAGACGGTGGAAAGACGCTCTTGATCACCCTCATCATGATGACAGATTTTGACAA 582
QY 41 ThrTyrGlnAlaIleLeuGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 593 ACCTATCAGGCACACATTTGGCATTGACCTCTTATCAAAAACAATGTACTTGGAGGATAGA 642
QY 61 ThrLeuGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
Db 643 ACCGTGCGATTGCAATTTATGGACACACAGCGTCAAGACGCGTTTCAGGAGCTTGATTCCT 702
QY 81 ArgTyrIleArgAspSerAlaAlaValValValValTyrAspIleThrAsnValAsnSer 100
Db 703 AGCTACATTCGAGACTCCACTGTGGCAGTTGTTGTTATGATATCACAATGTTAACTCA 762
QY 101 PheGlnGlnThrThrLysThrIleAspAspValArgThrGluArgGlySerAspValIle 120
Db 763 TTCCAGCAAACTACAAAATGATGATGATGTCAGAACAGAGAGAGGAAAGTGTGATCATC 822
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 823 ATCATGCTAGTAGGAAATAAAACAGACTCTGCTGATAAGAGCAAGTGTCAATTGAGGAG 882
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 883 GGAGAGAGGAAAGCCAAAGAGCTGAATGTTATGTTATTTGAACCCAGTGCAGAAAGCAGGA 942
QY 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 943 TACATGTGAAGCAGCTTTTCGGCGCTGTTGCTGCAGCTTTTACCTGGAATGGAAGCACA 1002
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 1003 CAGGACAGAAAGCAGAGAAAGACATGATTGACATATAAACTGGAAAGCCCTCAGGAGCAACCA 1062
QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
Db 1063 GTCAATGAAGCGCGCTGTTCTCTGCTTAATCCCCCTGGCATCTTTCAACCCCTCTCTGAGAAG 1122

```

```

Qy 221 -----ProPtyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeu 236
Db 1123 CTCAGTCTTTGGCCCACTACTCTTTCAATGACTGCAGTGTGAATATTGGCTTGAACCTT 1182
Qy 237 PheProSerLeuIleThrPheCysAsnSerSerLeuLeuProValSerTyrArg 254
Db 1183 TTCCCTTCAGTAATAACGATTTGCAATTCATATGCTGCCTGTTCGTGGAGG 1236

RESULT 5
BI694279
LOCUS 603347724F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5375477 5',
DEFINITION mRNA sequence.
ACCESSION BI694279
VERSION BI694279.1 GI:15656908
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11955 row: p column: 06
High quality sequence stop: 788.
FEATURES
    source
    1..790
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="FVB/N-3"
    /db_xref="taxon:10090"
    /clone="IMAGE:5375477"
    /tissue_type="tumor, biopsy sample"
    /dev_stage="5 months"
    /lab_host="DH10B"
    /clone_lib="NCI CGAP Mam2"
    /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
    Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
    Library constructed by Life Technologies. Investigator
    providing samples: Gilbert Smith, NIH"

ORIGIN
Alignment Scores:
Pred No.: 9,33e-111 Length: 790
Score: 1081.00 Matches: 217
Percent Similarity: 88.24% Conservative: 8
Best Local Similarity: 85.10% Mismatches: 25
Query Match: 82.90% Indels: 5
DB: 4 Gaps: 1

US-09-441-857-12 (1-254) x BI694279 (1-790)

Qy 5 GlyAspPheGlyAsnProLeuArgLysPheIleVal-PheLeuGlyGluInSerVa 24
Db 3 GGAGACTTCGGGAATCCGCTCAGGAAATTCAGCTGGTGGTTCCTGGGAGACGAGCGT 62
Qy 24 lAlaIysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsnThrTyrGlnAl 44
Db 63 TGGAAAGACGCTCTGTGATCCACCGCATTCATGTATGACAGATTTTGACACACTATCAGGC 122
Qy 44 aileileGlyIleAspPheLeuSerLysThrMetTyrLeuGluAspGlyThrIleGlyLe 64

```

```

Db 123 AACAAATTGGCATTGACCTTCTTATCAAAAACAATGCTACTTGGAGGATAGAACCGTGCATT 182
Qy 64 uArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIleProArgTyrIleAr 84
Db 183 GCAATTATGGGACACAGCAGGTCAAGAGCGGTTCAAGAGCTTGATTCTAGCTACATTCG 242
Qy 84 gApSerAlaAlaAlaValValTyrAspIleThrAsnValAsnSerPheGlnGlnTh 104
Db 243 AGACTCCACTGTGGCAGTTGTTTATGATATACAAATGTTAACTCATTTCCAGCAAC 302
Qy 104 rThrLysTyrIleAspAspValArgThrGluArgGlySerAspValIleIleThrLeuVa 124
Db 303 TACAAATGATTTGATGATGTCAGAACAGAGAGGGAAGTGTGTCATCATCATGTAGT 362
Qy 124 lGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGluGlyArgLys 144
Db 363 AGGAAATAAACAGATCTTGCTGATAGAGGCAAGTGTCAATTGAGGAGGAGAGGAA 422
Qy 144 sAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGlyTyrAsnVal 164
Db 423 AGCAAAAGAGCTGAATGTTATGTTTATGAACCAAGTGCAGGAGCAGATACAAATGAAA 482
Qy 164 sGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThrGlnAspGlySe 184
Db 483 GCAGCTTTTCGGCGTGTGCTGCAGCTTTACCTGGAATGGAAGCACACAGGACAGAAG 542
Qy 184 rArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThrValSerGluGl 204
Db 543 CAGAGAGACATGATTTGACATAAACTGGAAAAAGCCCTCAGGAGCAACAGTCAATGAAG 602
Qy 204 YGlyCysSerCysTyrSerProMetSerSerThrLeuProGlnLys----- 220
Db 603 CGGCTGTCTCTGCTAAATCCCTCGGCATCTTTCAACCTCTCGCAGAACTCAGTGT 662
Qy 221 -ProPtyrSerPheIleAspCysSerValAsnIleGlyLeuAsnLeuPheProSerLe 240
Db 663 GGCCCCATACTCTTTCAATGACTGCAGTGTGAATATTGGCTTGAACCTTTTCCCTT 722
Qy 240 uIleThrPheCysAsnSerSerLeuLeuProValSerTyrArg 254
Db 723 AATAACGATATTGCAATTCATCAATTGCTGCTGTTCGTGGAGG 765

RESULT 6
BI694279
LOCUS 7802212 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6042084
DEFINITION 5', mRNA sequence.
ACCESSION BI694279
VERSION BI694279.1 GI:21117590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 897)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13280 row: o column: 13
High quality sequence stop: 595.
FEATURES
    source
    1..897
    /organism="Homo sapiens"
    /mol_type="mRNA"

```

RESULT	7
CB320435	
LOCUS	744 bp mRNA linear EST 04-MAR-2003
DEFINITION	AGENCOURT_12397248 NIH_MGC_165 Mus musculus cDNA clone
ACCESSION	IMAGE:30282540 5', mRNA sequence.
VERSION	CB320435
KEYWORDS	EST.
SOURCE	CB320435.1 GI:28844670
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 744) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
TITLE	Contact: Robert Strausberg, Ph.D.
JOURNAL	Email: cgapbs-r@mail.nih.gov
COMMENT	Tissue Procurement: Dr. Leslie L. Heckert CDNA Library Preparation: CLONTECH Laboratories, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDCW43 row: 1 column: 13 High quality sequence stop: 539.

Alignment Scores:		
Pred. No.:	7.08e-109	Length:
Score:	1065.00	Matches: 897
Percent Similarity:	87.21%	Conservative: 218
Best Local Similarity:	84.50%	Mismatches: 7
Query Match:	81.67%	Indels: 27
DB:	5	Gaps: 6
		1

1	MetSerAlaGlyGlyAspPheGlyYanProLeuArgLysPheIysLeuValPheLeuGly	20
4	ATGTCACGGGCGGAGACTTCGGAAATCCGCTGAGGAAATTCAGCTGGTGTCTCTGGGG	63
21	GluInSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsn	40

50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Qy											Thr	Gly	Leu	Arg	Leu	Trp	Asp	Thr	Ala	Gly	Gln	Glu	Arg	Leu	Arg	Ser	Leu	Leu	Pro	
											:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db											Ac	Ag	Ta	Cg	Aat	Tg	Cga	Tat	Tgg	Gg	Cac	Ac	Gag	Cag	Tgt	Caa	Gcg	Ggt	Ttc	Agc

Qy 101 PheGlnGlnThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIle 120
|||
Db 304 TTCCAGCACAACCTACCAAGTGGATTGATGTCAGAACAGACAGAGGAAGTGAATGTTATC 363

Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db 424 GGAGAGAGAGAAAGCCAAAGAGCTGAATGTTATGTTATTTGAACTAGTGC AAAAGCTGA 483

		GlnAspGlySerArgGluaspMetSerAspileIylsLeuGluLysProGInGluThr	200
Qy			
		CAGGCAGAAGCAGAGAAGATATTGACATAAACTGGAAAAGCCCTCAGGAGCAACCA	603
Dδ			

QY	221	-----ProPrrTy-Ser-Phe-Ile-Asp-Cys-Ser-Val-Aen-Ile-Gly-Leu-Asn-Leu-	236
Dd	664	TCACTGCTTTGGCCCCCTTACTCTTCATTGAAGGTGCAGGGTGAAATATTGGCTTTGAACCTT	723

Qy	237	-PheProSerLeuIleThr-PheCysAanSerSerLeuLeuProValSer	252
		:	
		:	
		:	
Db	724	TTTCCTTCAATACGGAAATTGCATAATTTGCTGCCGTCT	773

Alignment Scores:

Alignment Scores:	
Pred. No.:	1,75e-103
Score:	1016.00
Percent Similarity:	88.33%
Best Local Similarity:	85.42%
Query Match:	77.91%
DB:	6
Gaps:	

US-09-441-857-12 (1-254) x CB320435 (1-744)

20 GlyGluGlnSerValAlaLyThrSerLeulleThrArgPheArgTyrAspSerPheAasp 39
 4 CGAGAGCAGAGCGTTGGAAAGACGCTCTTGATCACCCGATTTCATGTATGACAGTTTTGAC 63
 40 AsnThrTyrGlnAlaIleGlyLeuSerLybThrMetTyrLeuGluAasp 59

64 AACCCATATCAGGCACAAATGGCATTCGACTTCATTCAAAAACAAATGATCTTGGAGGAT 123

60 GlyThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIle 79

Db 124 AGAACCGTGGCGATTGCAATTATGGACACACACAGCGTCAAGAGCGGTCAGAGCGTTGATT 183

Qy 80 ProAAGTTrIIeAcGAspSerAlaAlaValValValTyrAspIleThrAsnValAsn 99

AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM14648 row: d column: 16
High quality sequence stop: 708.

FEATURES source

1. .797
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:5983585"
/tissue_type="cadpole"
/dev_stage="embryo, stages 40-45"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG XGC Emb8"
/note="Vector: pCIV-SF076.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.1 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores:
Pred. No.: 9,528-97 Length: 797
Score: 956.50 Matches: 193
Percent Similarity: 80.08% Conservative: 8
Best Local Similarity: 76.89% Mismatches: 29
Query Match: 73.35% Indels: 21
DB: 7 Gaps: 1
US-09-441-857-12 (1-254) x CFI51490 (1-797)

Qy 1 MetSerAlaGlyGlyAspPheGlyAenProLeuAtgLyPhelysLeuValPheLeuGly 20
Db 75 ATGTCCGGGGAGGAGACTTCGGGAACCGCTGAGGAATTCAACTTGTCTCTCTGGGG 134
Qy 21 GluGlnSerValAlaLysThrSerLeulleThrArgPheArgTyTyrAspSerPheAspAen 40
Db 135 GAGCAGAGCGTTGGGAAGACTTCTTTAATAACAGGTTTCATGTATGACAGCTTTGACAAAC 194
Qy 41 ThrTyGlnAlaIleleGlyleAspPheLeuSerLysThrMetTyriLeuGlyAspGly 60
Db 195 ACATACCAGCAACAACTTGAATTTGACTTCTTATCAAAACAAATGTACCTAGAGGATCGA 254
Qy 61 ThrileGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeullePro 80
Db 255 ACAGTCCGGTTGCAATTTGGGACACAGCAGGTTCAGGCGGTTTCAGAGATTGTATCCCT 314
Qy 81 ArgTyriLeuArgAspSerAlaAlaValValValTyriAspIleThrAsnValAenSer 100
Db 315 AGCTACATTCGGGACTCCACTGTGGCAGTTGTTTGTATATCACAATGTAAACTCA 374
Qy 101 PheGlnGlnThrThrystPheAspValArgThrGluArgLysSerAspValille 120
Db 375 TTCCAGCAACACCAAGAAATGGATCGATGACGTGCGAACAGACGGGGGACGACGTGATC 434
Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGly 140
Db 435 ATTATGTTAGTGGGAATTAACCGACCTGGCAGATAAAGACNAGTGTCCATTGAGGNA 494
Qy 141 GlyGluArgLysAlaLysGlyLeuAenValThrPheilleGluThrArgAlaLysThrGly 160

Db 495 GGAGAGCGGAAGCCAGGAGCTGAACGTAATGTTTATTGAAACGAGTGCAAGGCCGCGC 554
Qy 161 TyrAsnVallysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 555 TACATGTAAAGCAGCTCTTCGACCGCTGGCGCTGCCCTGGATGGAAAGCTCA 614
Qy 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGluGlnThr 200
Db 615 CAGGACAAGAGCAGAGAAGACATGATTGATCAAACTGGAAAGCCTCCAGAACAGCCA 674
Qy 201 ValSerGluGlyGlyCysSerCysTyriSerProMetSerSerSerThrLeuProGlnLys 220
Db 675 GTGTCCGAAGGGGCTGCTCTGTTATGCCCATCCCTCCCGGACTCCCCCATGAT 734
Qy 221 ProProTyriSerPheIleAspCysSerValAenIleGlyLeuAenLeuPheProSerLeu 240
Db 735 CTTCCC----- 740
Qy 241 IleThrPheCysAsnSerSerLeuLeuProVal 251
Db 741 -----TCCGGGACATCACTGCTCTCTGTC 764
RESULT 10
CA316338 792 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FW0-cbp-i-24-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
DEFINITION IMAGE:6812737 5', mRNA sequence.
ACCESSION CA316338
VERSION CA316338.1 GI:24534462
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 792)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES source

Location/Qualifiers
1. .792
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6812737"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FW0"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Ronald, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is ACCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN
Alignment Scores:
Pred. No.: 2,348-96 Length: 792
Score: 953.00 Matches: 194
Percent Similarity: 86.90% Conservative: 5
Best Local Similarity: 84.72% Mismatches: 22
Query Match: 73.08% Indels: 8
DB: 6 Gaps: 2

US-09-441-857-12 (1-254) x CA316338 (1-792)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLeuValPheLeuGly 20
Db 65 ATGTCGGCGGGGAGAGCTTCGGGAATCCGCTGAGGAAATTCAGCTGGTTCCTGGGA 124

Qy 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAsn 40
Db 125 GAGCAGAGCGTTGGAAAGACGCTTGTATCACCCGATTCATGTATGACAGTTTGCACAC 184

Qy 41 ThrTyrGlnAlaLeileGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 185 ACCTATCAGCAACATTTGGCATTCCTTATCAAAACATGATGATGAGGATAGA 244

Qy 61 ThrLeuGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuPro 80
Db 245 ACCGTGCGATTGCAATTTATGGACACAGCAGGTCAAGAGCGTTTCAGGAGCTTGATTCCT 304

Qy 81 ArgTyrLeuArgAspSerAlaAlaValValValTyrAspLeuThrAsnValAsnSer 100
Db 305 AGCTACATTCGAGACTCCACTGTGCGAGTGTGTTATGATATACAAAATGTTAACTCA 364

Qy 101 PheGlnGlnThrThrLysTrpLeuAspValArgThrGluArgGlySerAspValle 120
Db 365 TTCAGCAACACTACAAATGATGATGTCAGAACAGAGAGGAGGATGATGTCATC 424

Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db 425 ATCATGCTAGTAGGAATAAAACAGATCTTGTGATAAGAGCAAGTGTCTATTTGAGGAG 484

Qy 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheLeuGluThrArgAlaLysThrGly 160
Db 485 GGAGAGAGGAAGCCAAAGAGCTGAATGTTATGTTATTGAACCCAGTGCAAAAGCAGGA 544

Qy 161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
Db 545 TACATGTAAAGCAGCTTTTCGGGCTGTTGCTGAGCTTTACCTGGGAATGGAAGACACA 604

Qy 181 GlnAspGlySerArgGluAspMetSerAspLeuLysLeuGluLysProGlnGluGlnThr 200
Db 605 CAGGACAGAAGCAGAGAAGACATGATTGACATAAACTGGAAAGCCCTCAGGAGCAACCA 664

Qy 201 ValSerGluGlyGlyCysSerCysTyrSerPro-----MetSerSerSer 215
Db 665 GTCAATGAAGCGCGCTGTTCTCGCTTAATCCCTCGCATCTTCAACCCCTCGCAGAGCTCA 724

Qy 216 ThrLeuProGlnLysProProTyrSer 224
Db 725 CTGCTT-----GGCCCATCTCT 742

```

```

RESULT 11
LOCUS CB565915 628 bp mRNA linear EST 02-APR-2003
DEFINITION AGENCOURT 12691469 NIH MGC 146 Homo sapiens cDNA clone
IMAGE:6519185 5', mRNA sequence.
ACCESSION CB565915
VERSION CB565915.1 GI:29485445
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 628)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov

Tissue Procurement: Guthrie cDNA Resource Center
cDNA Library Preparation: Guthrie cDNA Resource Center
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: IRBF001 row: d column: 05
High quality sequence stop: 628.

Location/Qualifiers
1..628
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6519185"
/tissue_type="mixed"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_146"

FEATURES
SOURCE

/note="vector: pCDNA3.1; Site 1: multiple; Site 2: multiple; ORF's were PCR-amplified (from IMAGE clones or from commercially available cDNA libraries) and cloned by the Guthrie cDNA Resource Center (www.guthrie.org/cDNA) into pCDNA3.1. For specific information on cloning sites (which vary by clone), please refer to the Guthrie website, using the Guthrie ID given in the file ftp://image.llnl.gov/image.rearrayed_plates/IRBF.presv.dat
a. Note: this is a NIH_MGC Library."

ORIGIN

```

Alignment Scores:
Pred. No.: 1,71e-95 Length: 628
Score: 944.00 Matches: 188
Percent Similarity: 91.83% Conservative: 3
Best Local Similarity: 90.38% Mismatches: 17
Query Match: 72.39% Indels: 0
DB: 6 Gaps: 0

US-09-441-857-12 (1-254) x CB565915 (1-628)

Qy 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLeuValPheLeuGly 20
Db 1 ATGTCCAGCTGCGGAGAGCTTCGGGAATCCGCTGAGGAAATTCAGCTGGTTCCTGGGG 60

Qy 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyrAspSerPheAsn 40
Db 61 GAGCAAGCGTTGGAAAGACATCTTTCATCACCAGATTCATGATGACAGTTTTGCAAC 120

Qy 41 ThrTyrGlnAlaLeileGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
Db 121 ACCTATCAGGCAACATTTGGCATTCCTTATCAAAACATGATGATGAGGAGTCA 180

Qy 61 ThrLeuGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuPro 80
Db 181 ACAATCAGGCTTCAGCTGCGGATCTCGCGGTGAGGAACGTTCCGTAGCCTCATTTCCC 240

Qy 81 ArgTyrLeuArgAspSerAlaAlaValValValTyrAspLeuThrAsnValAsnSer 100
Db 241 AGTTACATCCGTGATTCGTCTGCGAGCTGTAGTAGTTACGATATACAAAATGTTAACTCA 300

Qy 101 PheGlnGlnThrThrLysTrpLeuAspValArgThrGluArgGlySerAspValle 120
Db 301 TTCAGCAAACTCAAAAGTGGATGATGATGTCAGAACACGAAAGAGAGGATGATGTTATC 360

Qy 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140

```

Db 361 ATCATGCTAGTAGGAATAAACAAGATCTTGTGACAGAGCAAGTGTCAATTGAGGAG 420
 QY 141 GYGLUArgLysGlyLeuAenValThrPheIleGluThrArgAlaLysThrGly 160
 Db 421 GGAGAGAGAGAAAGCCAAAGAGCTGAATGTTATTTGAACCTAGTGCAGAAAGCTGGA 480
 QY 161 TyrAenValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
 Db 481 TACAATGTAAGCAGCTCTTTCGACGTGTAGCAGCAGCTTTGCCGGAATGGAAGCACA 540
 QY 181 GlnAenGlySerArgGluAenMetSerAenPheLysLeuLysProGlnGluThr 200
 Db 541 CAGGACAGAGAGAGAGATATGATGACATAAACTGGAAGGCTCAGGAGCAACCA 600
 QY 201 ValSerGluGlyGlyCysSerCys 208
 Db 601 GTCAGTGAAGGAGGCTGTTCTCTGC 624

RESULT 12
 CA471443
 LOCUS
 DEFINITION AGENT0010698065 NCI CGAP 2K1d1 Danio rerio cDNA clone
 IMAGE:6791338 5', mRNA_sequence.

ACCESSION CA471443
 VERSION CA471443.1 GI:24927795
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov

Tissue procurement: Leonard I. Zon, M.D.
 cDNA Library Preparation: Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM14294 row: n column: 09
 High quality sequence stop: 721.

FEATURES

source
 1..870
 Location/Qualifiers
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:6791338"
 /lab_host="DH10B (T1-resistant)"
 /clone_lib="NCI CGAP 2K1d1"
 /note="Organ: kidney; Vector: pCMV-SPORT6.1; Site 1:
 EcoRV; Site 2: NotI; Cloned unidirectionally. Primer:
 Oligo dT. Average insert size 1.8 kb. Constructed by J.
 Wang (Research Genetics, Invitrogen Corp) from tissue
 donated by L. Zon (Harvard University). Note: this is a
 NCI CGAP Library."

ORIGIN

Alignment Scores:
 Pred. No.: 5,27e-95 Length: 870
 Score: 941.50 Matches: 193
 Percent Similarity: 79.84% Conservative: 13
 Best Local Similarity: 74.81% Mismatches: 35
 Query Match: 72.20% Indels: 17
 DB: 6 Gaps: 2

US-09-441-857-12 (1-254) x CA471443 (1-870)

QY 1 MetSerAlaGlyGlyAspPheGlyAenProLeuArgLysPheLysLeuValPheLeuGly 20

Db 87 ATGCTGCGAGGAGAGATTTCCGCAACCCGCTCCGCAAAATTCAGACTAGTTTTCTGGGA 146
 QY 21 GlnGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAen 40
 Db 147 GAGCAGACGCTGGGAAGACGCTGTTGATATACCAGATTATGTACGACAGTTTCGATAAC 206
 QY 41 ThrTyrGlnAlaIleIleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
 Db 207 ACCTACCAAGGCCACCATAGGAATTGATTTCTTGCGAAACCATGTATCTTGAGGACAGA 266
 QY 61 ThrIleGlyLeuArgLeuTyrAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
 Db 267 ACAGTGCAGCTTGGGACACAGCAGCAGGAGCGTTTCCGAGCTGTGATTTCCC 326
 QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAenValAenSer 100
 Db 327 AGCTACATCCGGGACTCGACGGTGTCTGTATGACATCACAATATGTCACATCA 386
 QY 101 PheGlnGlnThrThrLysTrpIleAspValArgThrGluArgGlySerAspValIle 120
 Db 387 TTCCAGCAGACCAACCAATGGATCGATCGATCCGACCGAAGAGAGGAGTGTATCATC 446
 QY 121 IleThrLeuValGlyAenArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
 Db 447 ATCATGCTAGTGGGAAACAAACACAGATCTTGACAGACAAAGCAAGTGTCTATCGAGAA 506
 QY 141 GlyGluArgLysAlaLysGlyLeuAenValThrPheIleGluThrArgAlaLysThrGly 160
 Db 507 GGAGAAAGGAAGCCAAAGAACTGAATGATTTATTTGAACCTAGTGTAAACCCGGA 566
 QY 161 TyrAenValLysGlnLeuPheArgValAlaAlaLeuProGlyMetGluSerThr 180
 Db 567 TACAACGTCAAGCAGCTTTTCCGCGGTGTAGCAGCAGCCCTTCTGTGTATGGAGAGCACA 626
 QY 181 GlnAspGlySerArgGluAspMetSerAspLysLeuGluLysProGlnGluThr 200
 Db 627 CAAGACAAGAGCAGAGAGACATGATGCACATAAAATTTGGAAAGCCACCAGAGCAGCG 686
 QY 201 ValSerGluGlyGlyCysSerCysTyrSerProMetSerSerSerThrLeuProGlnLys 220
 Db 687 GTCAGGAAGCGCGCTGTTTCATGCTGAGGC-----SerSerLeuLeuProValSerTrp 716
 QY 221 ProProTyrSerPheIleAspCysSerValAenIleGlyLeuAenLeuPheProSerLeu 240
 Db 717 -----TTCACTCGCTGTTTCTGTGCGCTGCTCTTACTGTGCTCTTTTCAGACGCTCTT 770
 QY 241 IleThrPheCysAen-----SerSerLeuLeuProValSerTrp 253
 Db 771 TATTCGCATCACAGCATCTCTCTGGATAATCTGCAGAAAGTCCGACCAACATGG 824

RESULT 13

CB067426

LOCUS

DEFINITION

KEYWORDS

ACCESSION

VERSION

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

UNPUBLISHED

CB067426

LOCUS

DEFINITION

KEYWORDS

ACCESSION

VERSION

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

UNPUBLISHED

CB067426 615 bp mRNA linear EST 21-JAN-2003
 IS36f10.x1 HR85 islet Homo sapiens cDNA clone IMAGE: 3', similar to
 SW:RAB6 HUMAN P20340 RAS-RELATED PROTEIN RAB-6. [1] ;, mRNA
 sequence.

CB067426
 CB067426.1 GI:27811946
 EST.
 SOURCE Homo sapiens (human)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 615)
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Teagareishvili,R.,
 Williams,T., Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium
 Unpublished (2000)

COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 444.

FEATURES

source

RESULT 14
CK696452
LOCUS
DEFINITION ZF101-P00073-DEPE-F2_F04 GISZF001_ra Danio rerio cDNA clone
IMAGE:7161798 5', mRNA sequence.
ACCESSION CK696452
VERSION
KEYWORDS
SOURCE
ORGANISM
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 788)
AUTHORS
Wei, C., Mathavan, S., Thoreau, H., Lim, L., Lee, C. and Ruan, Y.
TITLE
Genome Institute of Singapore, Zebrafish Gene Collection
JOURNAL
Unpublished (2004)
COMMENT
Contact: Ruan Y
Cloning and Sequencing
Genome Institute of Singapore
60 Biopolis Street, #02-01, Genome, Singapore 138672
Tel: +65 6478 8073
Fax: +65 6478 9059
Email: ruanyj@gis.a-star.edu.sg
GIS Clone ID: ZF101-P00073-BR2_F04
PCR Primers
FORWARD: M13
BACKWARD: M13
Plate: ZF101-P00073-BR2 row: F column: 04
Seq primer: CCGCATAACTTGATAGCA
High quality sequence stop: 788.

FEATURES

source

1..788
/organism="Danio rerio"
/mol_type="mRNA"
/strain="Singapore local strain"
/db_xref="taxon:7955"
/clone="IMAGE:7161798"
/tissue_type="Embryo"
/dev_stage="7 Different embryonic Stages(From just
fertilized Embryos to 72 hours just hatched baby fish)"
/lab_host="DH10B"
/clone_lib="GISZF001_ra"
/notes="Vector: pDNR-LIB; Site 1: Sfi A (GGCATTACGGCC);
Site 2: Sfi B (GGCGAGCGGCC); Priming method: Sfi-(dt)30
Primed; Priming sequence:
5.ATTCTAGGCGGAGCGCGGCACATG(T)30VN ; Directionally
cloned, 5' cloning site: Sfi A site GGCATTACGGCC ; 5'
linker/adaptor sequence: 5.AGCAGTGGTATCAACGAGATGGCC ;
3' cloning site: Sfi B site GGCGAGCGGCC ; 3'
linker/adaptor sequence: same as the priming sequence ;
Average insert size: 2kb ; For PCR insert analysis: Use
M13 Forward and reverse primers ; Library Amplified ;
Recombinants (inserts): 98% ; Library complexity: 5x10⁶ ;
Full-length construction (method): SMART, a Clontech
method The pooled tissue RNA was collected and used to
construct full length enriched cDNA library and also

ORIGIN

Alignment Scores:
Pred. No.: 1,32e-94 Length: 615
Score: 936.00 Matches: 186
Percent Similarity: 92.20% Conservative: 3
Best Local Similarity: 90.73% Mismatches: 16
Query Match: 71.78% Indels: 0
DB: 6 Gaps: 0
US-09-441-857-12 (1-254) x CB067426 (1-615)
QY 4 GlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGlyGluGlnSer 23
DB 1 GCGCGAGACTTCGGGAATCCGCTGAGGAAATTCAGCTGGTGTCTCGGGGAGCAAGC 60
QY 24 ValAlaLysThrSerLeuIleThrArgPheArgTyrAspSerPheAspAsnThrTyrGln 43
DB 61 GTTGAAGAAGACATCTTTGATCACCAGATTATGATATGACAGTTTGTGCAACACCTATCAG 120
QY 44 AlaIleLeuGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGlyThrIleGly 63
DB 121 GCAACAATGGCATTCATCTTTTATCAAAAACATATGATCTTGGAGGATCGAACATCAGG 180
QY 64 LeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIleProArgTyrIle 83
DB 181 CTTACGCTGTGGGATACTCGCGGTGAGGAACGTTTCCGTAGCCTCATTTCCAGTTACATC 240
QY 84 ArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSerPheGlnGln 103
DB 241 CGTGATTCCTGTCAGCTGTAGTAGTTTACATATACAAATGTTAACTCATTCATCCAGCA 300
QY 104 ThrThrLysTrpIleAspAspValArgThrGluArgGlySerAspValIleIleThrLeu 123
DB 301 ACTCAAAAGTGGATGATGATGTCAGACAGAGAGGAAGTGTATGTTATCATCATGCTA 360
QY 124 ValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlyGluArg 143
DB 361 GTAGGAATAAAACAGATCTTGCTGACAAAGAGCAAGTGTCAATTTGAGGGAGAGAGG 420

served as template to synthesize complex first strand cDNA probe. Two high density colony arrays were made from over 110K cDNA clones and hybridized with the probes. Low intensity clones were selected as they represented rare expressed clones. The hybridization intensities for all clones span from 0 to 1.8 million counts and the low abundant class ranged from 0 to 13,000."

ORIGIN

Alignment Scores:
Pred. No.: 1,898-94 Length: 788
Score: 936.00 Matches: 185
Percent Similarity: 90.87% Conservative: 4
Best Local Similarity: 88.94% Mismatches: 19
Query Match: 71.78% Indels: 0
DB: 7 Gaps: 0

US-09-441-857-12 (1-254) x CK696452 (1-788)

QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
DB 142 ATGTCTGCAGCAGGAGATTTCGGCAACCGCTCCGCAATTCAAGCTAGTTTCTGGGA 201
QY 21 GluGlnSerValAlaLysThrSerLeuLeuThrArgPheArgTyAspSerPheAspAsn 40
DB 202 GAGCAGAGCGTGGGAAGACGTCGTTGATTACAGATTATATACGACAGTTTCGATAAC 261
QY 41 ThrTyrGlnAlaIleTleGlyLeuAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
DB 262 ACCTACCAAGCCAGCATAGGAATTGATTTCTTGTGAAACCATGATTCGAGGACAGA 321
QY 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
DB 322 ACAGTCCGCTCGAGCTTTGGGACACAGCAGCAGGAGCGTTTCGCGAGCTGTATCCC 381
QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
DB 382 AGCTACATCCGGGACTCGACGGTGTCTGTGTAGTCTATGACATCACAATGTCAACTCA 441
QY 101 PheGlnGlnThrThrLysTrpIleAspValArgThrGluArgLysSerAspValIle 120
DB 442 TTCACGACAGACCACCAATGGATGATGATGATGATGATGATGATGATGATGATGATC 501
QY 121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
DB 502 ATCATGCTAGTGGGAACAAACAGATCTTGAGATAAAGGCAAGTGTCTATCGAGNA 561
QY 141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
DB 562 CGAGAAAGGAAAGCCAAAGAACTGAATGTAATGTTTATTGAACTAGTGTAAAGCCGA 621
QY 161 TyrAsnValLysGlnLeuPheArgValAlaAlaAlaLeuProGlyMetGluSerThr 180
DB 622 TACAACGTCACAGCAGCTTTCCGCGGTGAGCAGCAGCCCTTCCTGGTATGAGAGACA 681
QY 181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThr 200
DB 682 CAAGACAGAGCAGAGAGACATGATCGACATATAAATTGGAAGCCGCCAGCAGCCA 741
QY 201 ValSerGluGlyGlyCysSerCys 208
DB 742 GTCAGCGAAGCGCGTGTCTTTC 765

RESULT 15

BX852687
LOCUS BX852687 NICHED XGC_Emb4 Xenopus laevis mRNA linear EST 11-DEC-2003
DEFINITION ; IMAGE:5541856 5', mRNA sequence.
ACCESSION BX852687
VERSION BX852687.1 GI:39740496
KEYWORDS
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Xenopus.

REFERENCE

1 (bases 1 to 896)

AUTHORS

Heil, O., Neubert, P., Peters, M., Radelof, U., Schneider, D.,

Schroth, A., Korn, B. and Landgrebe, J.

Xenopus laevis UniGene Set 1 (RZPDLIB No.988)

TITLE

Unpublished (2003)

COMMENT

Contact: Ina Rolfes

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPDLIB; IMAGp998D1712240.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection (amp- resistant) (RZPDLIB

No.998) http://www.rzpd.de/cgi-

bin/products/showLib.pl.cgi/response?libNo=998 RZPDLIB; Xenopus

laevis UniGene Set 1 (RZPDLIB No.988) http://www.rzpd.de/cgi-

bin/products/showLib.pl.cgi/response?libNo=988 Contact: Ina Rolfes

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101

Fax: +49 30 32639 111

www.rzpd.de

This clone is available royalty-free from RZPD;

contact RZPD (clone@rzpd.de) for further information. Seq primer:

SP6, 5' ATTTAGGTGACACTATAG 3'.

Location/Qualifiers

1..696

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clones="IMAGp998D1712240 ; IMAGE:5541856"

/dev_stages="embryo, stage 31-32"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NICHED XGC_Emb4"

/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1;

NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dt. Average insert size 2.1 kb. Constructed by Life

Technologies. Note: This is a Xenopus Gene Collection

(XGC) library."

ORIGIN

Alignment Scores:

Pred. No.: 2,658-94 Length: 696
Score: 934.00 Matches: 184
Percent Similarity: 91.35% Conservative: 6
Best Local Similarity: 88.46% Mismatches: 18
Query Match: 71.63% Indels: 0
DB: 5 Gaps: 0

US-09-441-857-12 (1-254) x BX852687 (1-696)

QY 1 MetSerAlaGlyGlyAspPheGlyAsnProLeuArgLysPheLysLeuValPheLeuGly 20
DB 44 ATGTCGCGGAGAGAGACTTCGGGAACCCGCTGAGGAATTCAAACTGGTCTCTCTGGGG 103
QY 21 GluGlnSerValAlaLysThrSerLeuIleThrArgPheArgTyAspSerPheAspAsn 40
DB 104 GAACAGACGCTGGGAAGACATCTTATAACCAAGTTCATGATGACAGCTTTCACAAC 163
QY 41 ThrTyrGlnAlaIlelleGlyleAspPheLeuSerLysThrMetTyrLeuGluAspGly 60
DB 164 ACTTATCAGGCAACAATTTGGTATTGACTTCTTATCCAAACAATGTACCTAGAGATCGA 223
QY 61 ThrIleGlyLeuArgLeuTrpAspThrAlaGlyGlnGluArgLeuArgSerLeuIlePro 80
DB 224 ACAGTCCGGTTCGAATTTGTGGACACAGCAGGTTCAGGAGCTTCAGGAGTCTGATCCCT 283
QY 81 ArgTyrIleArgAspSerAlaAlaValValValTyrAspIleThrAsnValAsnSer 100
DB 284 AGCTACATCCGAGCTCCACTGTCGACGTTGTTGTTTGTATATACAAATGTAACCTCA 343
QY 101 PheGlnGlnThrThrLysTrpIleAspValArgThrGluArgLysSerAspValIle 120

```
Db      344 TTCCAGCAAAACACAAATGGATCGATGATGTCGGAACAGAGCGGGGAGTGATGTGATC 403
Qy      121 IleThrLeuValGlyAsnArgThrAspLeuAlaAspLysArgGlnValSerValGluGlu 140
Db      404 ATTATGCTTGTGGGAAATAAAACAGATCTTGGGATAAAAGGCAAGTGTCATTGAGGAA 463
Qy      141 GlyGluArgLysAlaLysGlyLeuAsnValThrPheIleGluThrArgAlaLysThrGly 160
Db      464 GGGGAGCGGAAAGCCAGAGCTGAACGTAATGTTATTGAAACCAAGTGCAAGGCTGGC 523
Qy      161 TyrAsnValLysGlnLeuPheArgArgValAlaAlaLeuProGlyMetGluSerThr 180
Db      524 TACAATGTAAGCAGCTCTTCGACGTGTTGCAGCGCGCTTGCCCTGGGATGGAAAGCTCA 583
Qy      181 GlnAspGlySerArgGluAspMetSerAspIleLysLeuGluLysProGlnGlnThr 200
Db      584 CAGGATAAGAGCAGAGAGACATGATTGATATCAAACTGGAAAGCCCTCCAGAACAGCCA 643
Qy      201 ValSerGluGlyGlyCysSerCys 208
Db      644 GTGTCGGAAGGAGGCTGCTCCTGT 667
```

Search completed: April 25, 2005, 05:40:55
Job time : 3110 secs